Structure and function of class I diterpene synthase

Ke Zhou
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd
Part of the Biochemistry Commons

Recommended Citation
https://lib.dr.iastate.edu/etd/13223

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
# Table of Contents

List of Abbreviations Commonly Used in This Thesis iv

## Chapter I: General introduction

1. Terpenoids biosynthesis 1
2. Terpene synthase (Terpene cyclase) 2
3. Labdane-related diterpenes in crop plant 6
4. Plasticity of class I terpene synthase 7
5. References 10
6. Figures 15

## Chapter II: Insights into diterpene cyclization from the structure of the bifunctional abietadiene synthase from *Abies grandis* 20

1. Abstract 20
2. Introduction 20
3. Results 23
4. Discussion 27
5. Methods 28
6. Acknowledgement 30
7. References 30
8. Figures 35

## Chapter III: Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants 41

1. Abstract 41
2. Introduction 42
3. Results 44
4. Discussion 45
5. Concluding remarks 47
6. Experimental 47
7. Acknowledgement 50
8. References 50
9. Figures 53

## Chapter IV: Insights into the evolution of diterpenoid metabolism in cereal crop plants from functional characterization of wheat diterpene synthases 57

1. Abstract 57
2. Introduction 58
3. Results 59
4. Discussion 63
5. Conclusion 66
6. Experimental 66
7. Acknowledgement 69
<table>
<thead>
<tr>
<th>Chapter V: Direct action of a single residue switch for diterpene synthase product outcome</th>
<th>81</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>81</td>
</tr>
<tr>
<td>Introduction</td>
<td>81</td>
</tr>
<tr>
<td>Results</td>
<td>83</td>
</tr>
<tr>
<td>Discussion</td>
<td>84</td>
</tr>
<tr>
<td>Conclusion</td>
<td>86</td>
</tr>
<tr>
<td>Experimental</td>
<td>87</td>
</tr>
<tr>
<td>References</td>
<td>89</td>
</tr>
<tr>
<td>Figures</td>
<td>92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter VI: Electrostatic effects on (di)terpene synthase product outcome</th>
<th>98</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>98</td>
</tr>
<tr>
<td>Introduction</td>
<td>99</td>
</tr>
<tr>
<td>Terpene synthases</td>
<td>100</td>
</tr>
<tr>
<td>Basic catalytic mechanism</td>
<td>100</td>
</tr>
<tr>
<td>Substrate folding – the template enzyme model</td>
<td>101</td>
</tr>
<tr>
<td>Kinetic rather than thermodynamic control</td>
<td>101</td>
</tr>
<tr>
<td>A role for the pyrophosphate anion co-product</td>
<td>103</td>
</tr>
<tr>
<td>A single residue switch for diterpene synthase product outcome</td>
<td>104</td>
</tr>
<tr>
<td>Labdane-related diterpene synthases</td>
<td>105</td>
</tr>
<tr>
<td>Discovery</td>
<td>105</td>
</tr>
<tr>
<td>Extension</td>
<td>108</td>
</tr>
<tr>
<td>Application</td>
<td>109</td>
</tr>
<tr>
<td>Stereospecificity</td>
<td>111</td>
</tr>
<tr>
<td>Conclusions</td>
<td>112</td>
</tr>
<tr>
<td>References</td>
<td>113</td>
</tr>
<tr>
<td>Figures</td>
<td>117</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter VII: A secondary position involved in catalytic specificity of labdane-related diterpene synthase</th>
<th>121</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>121</td>
</tr>
<tr>
<td>Introduction</td>
<td>121</td>
</tr>
<tr>
<td>Results</td>
<td>123</td>
</tr>
<tr>
<td>Discussion</td>
<td>124</td>
</tr>
<tr>
<td>Conclusion</td>
<td>126</td>
</tr>
<tr>
<td>Experimental</td>
<td>126</td>
</tr>
<tr>
<td>References</td>
<td>128</td>
</tr>
<tr>
<td>Figures</td>
<td>130</td>
</tr>
</tbody>
</table>

| Chapter VI: Conclusions and Future directions                                                           | 131 |
| References                                                                                             | 136 |
**List of Abbreviations Commonly Used in This Thesis:**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AgAS</td>
<td><em>Abies grandis</em> abietadiene synthase</td>
</tr>
<tr>
<td>AtKS</td>
<td><em>Arabidopsis thaliana</em> kaurene synthase</td>
</tr>
<tr>
<td>CPP</td>
<td>copalyl diphosphate</td>
</tr>
<tr>
<td>CPS</td>
<td>copalyl diphosphate synthase</td>
</tr>
<tr>
<td>diTPS</td>
<td>diterpene synthase</td>
</tr>
<tr>
<td>DMAPP</td>
<td>dimethylallyl diphosphate</td>
</tr>
<tr>
<td>FPP</td>
<td>farnesyl diphosphate</td>
</tr>
<tr>
<td>GA</td>
<td>gibberellic acid</td>
</tr>
<tr>
<td>GC</td>
<td>gas chromatography</td>
</tr>
<tr>
<td>GC-FID</td>
<td>gas chromatography – flame ionization detection</td>
</tr>
<tr>
<td>GC-MS</td>
<td>gas chromatography – mass spectrometry</td>
</tr>
<tr>
<td>GPP</td>
<td>geranyldiphosphate</td>
</tr>
<tr>
<td>GGPP</td>
<td>geranylgeranyldiphosphate</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>HvKS</td>
<td>barley (<em>Hordeum vulgare</em>) kaurene synthase</td>
</tr>
<tr>
<td>IPP</td>
<td>isopentenyl pyrophosphate</td>
</tr>
<tr>
<td>KS</td>
<td>kaurene synthase</td>
</tr>
<tr>
<td>KSL</td>
<td>Kaurene synthase-like</td>
</tr>
<tr>
<td>MEP</td>
<td>2-Methyl-D-erythritol-4-phosphate pathway</td>
</tr>
<tr>
<td>MEV</td>
<td>mevalonic acid synthesis pathway</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>OsKS(L)</td>
<td>rice (<em>Oryza saliva</em>) kaurene synthase (-like)</td>
</tr>
<tr>
<td>SAD</td>
<td>single anomalous dispersion</td>
</tr>
<tr>
<td>TaKS(L)</td>
<td>wheat (<em>Triticum aestivum</em>) kaurene synthase (-like)</td>
</tr>
<tr>
<td>TPS</td>
<td>terpene synthase</td>
</tr>
</tbody>
</table>
Chapter I: General introduction

Terpenoids are the largest class among all known natural products, with over 55,000 known members. These compounds are particularly widespread in plants, where they are involved in primary metabolism to adjust plant growth and development, like gibberellins phytohormones, but also act as secondary metabolites playing roles in the interactions of plants with their environment, such as plant pollination, defense, and signaling. Notably, some terpenoids already have demonstrated usefulness as biomaterials and pharmaceuticals. For example, Taxol, a diterpenoid, is a potent anticancer drug that was isolated and identified from the bark of the Pacific yew (Arbuck & Blaylock 1995; Wani et al. 1971). The labdane-related diterpenoids are a special group, consisting of over 7,000 members, which is distinguished by their unique biosynthesis (Peters 2010). Gibberellins phytohormones as well as antibiotics such as some of phytoalexins and phytoanticipins fall into this family (VanEtten et al. 1994).

Terpenoids biosynthesis

Terpenoids are biosynthesized by multiple complex steps. Firstly, isoprenyl diphosphate synthases (also named prenyltransferases) catalyze condensation reaction from the universal C5 precursors, dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) (Ruzicka 1994), which originate from two pathways in plants: the mevalonate (MEV) pathway present in the cytosol/endoplasmic reticulum, and the 2C-methyl-D-erythritol-4-phosphate (MEP) pathway occurring in the plastids (Rohmer et al. 2003), and produce linear isoprenoids of incremental length, including geranyl diphosphate (GPP, 10C), farnesyl diphosphate (FPP, 15C), geranylgeranyl diphosphate (GGPP, 20C). Then corresponding monoterpene, sesquiterpene and diterpene synthases catalyze cyclization reaction in which
terpenes with remarkably diverse hydrocarbon backbones are generated. The bioactive natural products, terpenoids, are produced by subsequent elaboration wherein the hydrocarbon backbones are decorated with oxygen by cytochromes P450. The studies this dissertation will largely involve are relevant to the middle step: cyclization reactions in labdane-related diterpenoid biosynthesis. Generally, such labdane-related diterpenoids, just like gibberellin biosynthesis, are derived from characteristic labdadienyl/copalyl diphosphate (CPP) biosynthetic intermediates, which containing a fused bicyclic core, from which the ‘labdane-related diterpenoid’ nomenclature was originated (Peters 2010). In particular, this is produced by bicyclization of the universal diterpenoid precursor \((E,E,E)\)-geranylgeranyl diphosphate (GGPP) catalyzed in protonation-initiated reactions by class II diterpene synthase to form the characteristic CPP with normal, \(ent\)- or \(syn\)- stereochemistry. Then, this bicyclic diphosphate intermediate is further cyclized and/or rearranged by class I diterpene synthases catalyzing ionization-initiated reactions (Davis and Croteau, 2000), to generate the final labdane-related diterpene olefin.

**Terpene synthase (Terpene cyclase)**

During the multiple steps in production of active terpenoid natural products, terpene synthases (TPS) are a critical group of enzymes because they create diverse hydrocarbon skeletal/backbone structures by catalyzing complex cyclization and/or rearrangement, the most complex chemical reactions occurring in nature, from linear isoprenyl pyrophosphate precursors (Christianson 2006). Thus, their products underlie the amazing variety of terpenoid natural products. Interestingly, all x-ray crystallographic studies on TPS from lower microbe to higher plant over 10 years revealed that they share a homologous fold to achieve such extensive product diversity with structural and stereochemical precision, despite
of absence of high sequence identity (Starks et al. 1997; Lesburg et al. 1997; Rynkiewicz et al. 2001; Whittington et al. 2002) (Fig. 1). Actually, there is less than 15% primary sequence identity among plant, fungal and bacterial TPS.

According to these structures, TPS generally adopts two kinds of folds (Fig. 2). Typically, the ionization-initiated TPS have been designated class I terpene synthases. Its active site is located within an α domain, which adopts a common α-bundle fold wherein a canonical aspartate-rich DDXX(XX)D/E motif along with a secondary and less well conserved (N,D)DXX(S,T,G)XXXE motif together bind a magnesium cluster that triggers the departure of the substrate diphosphate leaving group, and concurrently initiates the cyclization and rearrangement reaction. In the middle position of the secondary metal binding motif, a glycine is frequently found in place of the prototypical Ser/Thr side chain hydroxyl group. How this can be involved in ligating one of the three catalytically requisite divalent metal ions will be discussed at one chapter of this dissertation (Zhou & Peters 2009). In contrast, the protonation-initiated TPS have been designated class II terpene synthases. The corresponding active sites resides between β/γ domains, both of which exhibit a α-barrel fold in which a DXDD motif in the β domain provides the proton donor that triggers initial carbocation formation (Christianson 2006). The γ fold exhibits a similar topology with β fold (Wendt et al. 1997; Köksal et al. 2011).

Unlike microbial class I TPS displaying only α fold (Lesburg et al. 1997), most plant monoterpene and sesquiterpene synthases adopt αβ assembly in which the α fold executes its typical function, but β fold is inactive due to the absence of characteristic DXDD motif (Starks et al. 1997; Whittington et al. 2002). Microbial class II diterpene synthases (diTPS) display βγ domain structure that resembles that of triterpene synthases (Squalene-hopene
cyclase, SHC) (Wendt et al. 1997). Interestingly, both phylogenetic and bioinformatic analysis suggest that most plant diTPS display a composite domain structure of γβα (Cao et al. 2010; Trapp & Croteau 2001; Bohlmann et al. 1998), in which the N-terminus γ domain with ~240-residues is usually called an “insertional” domain, regardless of function (i.e., class I/class II bifunction, or either single function; Fig. 3). The recently determined structures of the class I diTPS, taxadiene synthase from Taxus brevifolia (TbTS), the class II diTPS copalyl diphosphate synthase from Arabidopsis thaliana (AtCPS), as well as the bifunctional (class II/I) abietadiene synthase from Abies grandis (AgAS), uncovered this kind of assembly for plant diTPS (Köksal et al. 2011a; Köksal et al. 2011b; Zhou et al. manuscript preparation). However, exceptions lacking insertional domain were also found among class I diTPS, such as casbene synthase from castor bean, OsKSL2 from rice and DsKSL in Danshen (Mau & West 1994; Xu et al. 2007; Gao et al. 2009). All these diTPS consist of βα domains like mono- and sesqui-TPS. In contrast, very few mono- or sesqui-TPS contain the insertional element (corresponding to the γ domain), with the (E)-α-bisabolene synthase in grand fir and linalool synthase in Clarkia among the only known examples (Bohlmann et al. 1998; Dudareva et al. 1996).

It has been around 10 years since the first structures of most groups of TPS have been solved, which included mono-, sesqui-, tri- TPS and isoprenyl diphosphate synthases (Starks et al. 1997; Lesburg et al. 1997; Whittington et al. 2002; Wendt et al. 1997; Tarshis et al. 1994). But until this last year, no crystal structure of a diterpene synthase had yet been reported. In this dissertation, the insights into diterpene cyclization from the structure of abietadiene synthase (AgAS) will be discussed.
Grand fir (*Abies grandis*) produces the oleoresin which plays role in forming a physical barrier that seals the wound after bark beetles attack it. (Johnson & Croteau 1987). The oleoresin contains resin acids derived largely from the abietane family of diterpene olefins (Funk et al. 1994). Abietadiene synthase (AS) of grand fir catalyzes the cyclization and rearrangement of the GGPP to a mixture of abietadiene isomers, representing the committed step in resin acid biosynthesis (Vogel 1996). Specifically, AgAS is a bifunctional enzyme that catalyzes two sequential cyclization reactions in distinct active sites (Peters et al. 2000). Firstly, in the class II active site, protonation at the C14 atom of GGPP triggers a series of electrophilic attacks on carbocations by double bonds and deprotonation at C19 to produce a stable bicyclic intermediate copalyl diphosphate (normal CPP) (Ravn 1998). Then this intermediate undergoes active site shift by diffusion to the active site of class I domain (Peters 2002). There ionization of the allylic diphosphate ester triggers further cyclization, intramolecular proton transfer, methyl migration and deprotonation at different positions to yield abietadiene isomers (Ravn 2000; Peters 2000; Ravn 2002, Fig. 4). Thus, AgAS is a bifunctional labdane-related diterpene synthase.

**Labdane-related diterpenes in crop plant**

Wheat (*Triticum aestivum*) and rice (*Oryza sativa*) are two of the most agriculturally important cereal crop plants, and together with maize, provide 60 percent of the world's food energy intake. Recently, enabled by rice draft genome sequences (Goff et al., 2002; Yu et al., 2002), biosynthesis of natural products in rice metabolism have been widely studied. Particularly that of the labdane-related diterpenoids, with all the functional diTPS having been characterized. In detail, rice contains a family of ent-kaurene synthase like (KSL) genes, which encode a set of diTPS involved in the second (ionization-initiated) cyclization step,
forming the various hydrocarbon backbones required for production of the observed natural products (Xu et al., 2007). This includes the *ent*-kaurene synthase (KS) required for gibberellin phytohormone biosynthesis for plant growth and development (Margis-Pinheiro et al., 2005; Sakamoto et al., 2004), as well as the KSL involved in phytoalexins and allelochemical biosynthesis (Swaminathan et al., 2009; Wang et al., 2011; Kato-Noguchi and Ino, 2003; Kato-Nogu-chi et al., 2002) (Fig. 5). Wheat (*Triticum aestivum*) is an important cereal crop, and closely related to rice. It contains a similarly expansive family of KSL genes (*TaKSL*) as well. In addition, from another small grain cereal barley (*Hordeum vulgare L.*), an *HvKS* gene was isolated (Spielmeyer et al., 2004). One chapter of this dissertation will present isolation, cloning and biochemical characterization of wheat KSL genes, as well as demonstration of enzymatic function of *HvKS*, along with general analysis of these genes and their products, to provide insight into the evolution of diterpenoid metabolism, at least in the cereal crop plant family.

**Plasticity of class I terpene synthase**

Recent studies have demonstrated that terpene synthases exhibit extreme plasticity, with small numbers of amino acid residue substitutions being sufficient to substantially alter product outcome (Xu et al. 2007; Wilderman et al. 2007; Morrone et al. 2008; Keeling et al. 2008; Yoshikuni et al. 2006; Kampranis et al. 2007). Typically, these plastic residues reside in the active pocket. For instance, similar but subtly different metal binding interactions among class I TPS are considered important in stabilizing the proper active site contour and consequently are a possible strategy for divergent product outcomes. That is confirmed by the interactions between substrate pyrophosphate and trinuclear Mg$^{2+}$ cluster liganded by the classic DDXX(D/E) motif and the secondary, (N/D)DX$_2$(S/T)X$_3$E, metal binding motif, as
demonstrated by the structures of trichodiene, (+)-bornyl diphosphate and epi-aristolochene synthases (Starks et al. 1997; Rynkiewicz et al. 2001; Whittington et al. 2002), and by mutational analysis of the aspartate-rich motif of trichodiene synthase (Cane et al. 1996). Besides these catalytic motifs, other important residues relevant to product profile are located at active site. For example, two active site regions, a conserved kink structural feature and an Asn side chain, were identified as being responsible for the product specificity of the 1,8-cineole synthase from Salvia fruticosa (Kampranis et al. 2007). Another successful example is discovery of seven new sesquiterpene synthases with specific reaction pathways to produce very different products by performing exhaustive site-directed mutagenesis at 3-5 residues among 19 residues surround the active site of a promiscuous γ-humulene synthase (Yoshikuni et al. 2006). However, it must be noted that surrounding layers enclosing the active site also offer contribution to the product specificity. Such relevance has already been demonstrated for TEAS and HVS (Greenhagen et al. 2006), which contain the same residues in their active sites, but yield differing product profiles. All these mutagenesis studies, as well as reported crystal structures, support a template model that implies TPS dictate reaction outcome by restricting their substrate and discrete reaction intermediates to a subset of the possible conformations (Christianson 2008).

Among the most interesting recent results are those from our group demonstrating the ability of single residue changes to dramatically shift the product outcome of terpene synthases involved in labdane-related diterpene biosynthesis, which highlight the importance of electrostatic effects exerted by TPS (Xu et al. 2007; Wilderman et al. 2007; Morrone et al. 2008; Zhou & Peters 2011). This was originally based on the identification of highly homologous but functionally divergent paralogs from rice (Oryza sativa). Specifically,
OsKSL5i and OsKSL5j, which are 98% identical at amino acid level, with modeled enzyme structures indicating that there are only three differences in the active site, yet which produce ent-(iso)kaur-15-ene or ent-pimaradiene, respectively. This can be rationalized by ionization-initiated cyclization of ent-CPP to a pimaren-8-yl+ intermediate that may be followed by either direct deprotonation to ent-pimaradiene, as catalyzed by OsKSL5j, or secondary cyclization to a beyeran-16-yl+ intermediate that undergoes ring rearrangement to the kauranyln ring structure prior to deprotonation, as catalyzed by OsKSL5i (Xu et al. 2007). Thr substitution for an active site Ile was found to be sufficient to 'short circuit' the complex cyclization and rearrangement reaction catalyzed by OsKSL5j, leading to the production of an ent-pimaradiene instead (Xu et al. 2007). Strikingly, this position is conserved as Ile in the ent-kaur-16-ene synthases found in all higher plants for gibberellin phytohormone biosynthesis and, despite sharing only 41–52% aa sequence identity with the rice ent-isokaurene synthases, Thr substitution had a very similar effect in two disparate ent-kaurene synthases, “short-circuiting” the more complex cyclization of ent-CPP to ent-kaurene to production of the “simpler” ent-pimaradiene. In addition, the effect on product profile of this single switch has been extended by investigations with normal and syn-CPP specific diTPS (Wilderman et al. 2007; Morrone et al. 2008). In particular, Ile substitution for the corresponding Thr in a syn-pimaradiene synthase was found to dramatically increase the complexity of the catalyzed reaction, resulting in predominant production of a rearranged tetracyclic aphidicolene rather than tricyclic pimaradiene (Morrone et al. 2008). Conversely, substitution of Ala by Ser in the normal CPP specific abietadiene synthase (AgAS) results in a diTPS that produces largely pimaradienes (Wilderman et al. 2007). Regardless of CPP stereochemistry, it has been hypothesized that the relevant residue directly interacts with the
initial carbocation in the pimarenyl+ intermediate formed by initial cyclization of the relevant CPP (Xu et al. 2007; Wilderman et al. 2007; Morrone et al. 2008). Specifically, it is hypothesized that the side chain of this product switch residue is adjacent to the pimar-15-en-8-yl carbocation, such that the presence of a hydroxyl dipole stabilizes this intermediate long enough for deprotonation to occur, while the lack of such stabilization (i.e. in the presence of an aliphatic residue) enables the tightly bound pyrophosphate anion co-product to steer carbocation migration towards itself, followed by subsequent rearrangement to a more stable tertiary carbocation, which is still adjacent to the pyrophosphate counter-ion, prior to concluding deprotonation. A further study that this residue switch for product outcome in labdane-related diTPS is specific with respect to both active site location and side chain chemical composition will be presented in this dissertation, supporting a direct action of this single residue switch for labdane-related diTPS product outcome.

Recent studies also revealed plasticity of TPS on substrate recognition. For instance, most plant class I labdane-related diTPS react specifically with single stereoisomers of CPP (Xu et al., 2007; Peters et al., 2000; Yamaguchi et al., 1998). Interestingly, several identified rice and wheat kaurene synthase-like enzymes exhibit dual substrate plasticity, and can utilize two stereoisomers of CPP to produce different diterpene skeletons with similar levels of activity (Morrone et al., 2011). Specifically, OsKSL4 and OsKSL11 react with syn- and normal CPP even though rice doesn’t contain a normal CPP synthase. By contrast, wheat makes normal and ent-CPP, but does not seem to make syn-CPP (Wu et al., submitted; Toyomasu et al., 2009). However, besides using CPP of normal stereochemistry, the ability of TaKSL 1 and TaKSL 4 to react with syn-CPP suggests that by means of enzymatic
plasticity, rapid evolution could occur in cereal crop diterpenoid metabolism, leading to the
generation of new natural products.

References


conifers: cDNA cloning, characterization, and functional expression of wound-
inducible (E)-alpha-bisabolene synthase from grand fir (Abies grandis). Proc Natl
Acad Sci U S A. 95(12):6756-61.

Bohlmann, J., Meyer-Gauen, G., & Croteau, R., 1998 Plant terpenoid synthases:
Molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. USA 95:4126-
4133.

Cane, D.E., Xue, Q., Fitzsimons, B.C., 1996. Trichodiene synthase. Probing the role of the
highly conserved aspartate-rich region by site-directed mutagenesis. Biochemistry.
35(38):12369-76.

Cao, R., Zhang, Y., Mann, F.M., Huang, C., Mukkamala, D., Hudock, M.P., Mead,
M.E., Prisic, S., Wang, K., Lin, F.Y., Chang, T.K., Peters, R.J., and Oldfield, E.,

Christianson, D.W., 2006 Structural biology and chemistry of the terpenoid cyclases. Chem.
Rev. 106(8): 3412-3442


monoterpenes, sesquiterpenes, and diterpenes. Top. Curr. Chem. 209: 53-95

Dudareva, N., Cseke, L., Blanc, V.M., Pichersky, E., 1996. Evolution of floral scent in
Clarkia: novel patterns of S-linalool synthase gene expression in the C. breweri

Funk, C., Croteau, R., 1994. Diterpenoid resin acid biosynthesis in conifers:
characterization of two cytochrome P450-dependent monooxygenases and an
aldehyde dehydrogenase involved in abietic acid biosynthesis. Arch Biochem

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J.,
Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C.,
Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J.,
Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper,
B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y.,
Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A.,
Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller,
R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus,
genome (Oryza sativa L. ssp. japonica). Science 296, 92–100.


Zhou, K., Mann, F. M., Hoy, J. A., Honzatko, R. B. and Peters, R. J., (Manuscript preparation) Insights into diterpene cyclization from the structure of the bifunctional abietadiene synthase from Abies grandis

**Figures**

Figure 1: Superimposed crystal structures of 5-epi-aristolochene synthase from tobacco (green), Bornyl Pyrophosphate synthase from Salvia officinalis (Cyan), and Pentalenene Synthase from *Streptomyces sp. UC5319* (Magentas) show homologous fold of class I Terpene Synthase (Starks et al. 1997; Lesburg et al. 1997; Whittington et al. 2002).
Figure 2: A: Pentalenene Synthase from *Streptomyces sp* shows a typical $\alpha$-bundle fold (Lesburg et al. 1997), while Squalene-hopene cyclase from Alicyclobacillus acidocaldarius displays $\alpha$-barrel fold (Wendt et al. 1997).
Figure 3: Schematic diagram of terpene synthases illustrating general structural features and locations of the catalytic aspartate-rich elements.

<table>
<thead>
<tr>
<th>Microbial class I TPS</th>
<th>DDXXD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant class I TPS</td>
<td>DDXXD</td>
</tr>
<tr>
<td>(Most Monoterpene &amp; Sesquiterpene Synthases)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Class I Domain</td>
</tr>
<tr>
<td>Microbial class II TPS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXDD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant class II TPS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXDD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Plant class I TPS</td>
<td></td>
</tr>
<tr>
<td>(Some of diTPS)</td>
<td>DDXXD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insertional Domain</td>
</tr>
<tr>
<td></td>
<td>Class II Domain</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Special class I and II TPS</td>
<td></td>
</tr>
<tr>
<td>(Bifunctional diTPS in lower plant)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXDD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DDXXD</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insertional Domain</td>
</tr>
<tr>
<td></td>
<td>Class II Domain</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Class I Domain</td>
</tr>
</tbody>
</table>
Figure 4: Sequential cyclization reactions catalyzed by AgAS. (Peters et al. 2000)
Figure 5: Known labdane-related diterpene cyclization reactions in rice. The corresponding cyclases are indicated, along with their products and, where known, the derived natural products (dashed arrows indicate multiple biosynthetic steps). (Xu et al. 2007)
Chapter II: Insights into diterpene cyclization from the structure of the bifunctional abietadiene synthase from *Abies grandis*

Manuscript in preparation and to be submitted to Journal of Biological Chemistry

Ke Zhou, Julie A. Hoy, Francis M. Mann, Richard B. Honzatko, Reuben J. Peters*

Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

Abstract

Abietadiene synthase from *Abies grandis* (AgAS) has served as a model for investigation of diterpene synthase activity, and here we report its crystal structure at 2.3 Å resolution. This bifunctional enzyme catalyzes both class I (ionization-initiated), as well as class II (protonation initiated) cyclization reactions, and is composed of three α-helical domains, designated α, β, and γ, wherein the class I active site falls within the C-terminal α domain and the class II active site lies between the N-terminal γβ domains. The structure not only clarifies the evolutionary origins of diterpene synthases, but also provides insights into the enzymatic structure-function relationships underlying both catalyzed reactions, as well as a recently discovered regulatory mechanism.

Introduction

Terpenoids form the largest group of natural products, with some 50,000 known\(^1\). Underlying the observed diversity are the manifold hydrocarbon skeletal backbones produced by the cyclization and/or rearrangement of acyclic precursors catalyzed by the relevant terpene synthases\(^2\). These are first sub-divided based on the number of five-carbon isoprenoid repeats; with the ten-carbon monoterpenoids generally derived from geranyl diphosphate, the fifteen-carbon sesquiterpenoids from farnesyl diphosphate, the twenty-
carbon diterpenoids from geranylgeranyl diphosphate (GGPP), and the thirty-carbon triterpenoids from squalene.

As might be suspected from the composition of their precursors, the skeletal backbone structures of the lower (mono-, sesqui-, and di-) terpenes are generally produced by cyclization and/or rearrangement reactions initiated by ionization of the substrate allylic diphosphate ester bond, while triterpenes are cyclized via a protonation-initiated reaction (i.e., of the terminal C=C double bond or oxido-ring of the derived oxido-squalene). The corresponding enzymes have been termed class I and class II, respectively. Notably, biosynthesis of the large super-family of labdane-related diterpenoids (~7,000 known) combines both types of reactions, in which GGPP is bicyclized by a class II enzyme prior to further cyclization and/or rearrangement catalyzed by a class I enzyme.

The labdane-related diterpenoids include many important plant natural products, such as the gibberellin phytohormones and others that act in defense (e.g., the ubiquitous conifer resin acids). Indeed, it has been suggested that the terpene synthases involved in gibberellin biosynthesis are the ancestral progenitors to the extensive family of lower (i.e., class I) terpene synthases found in plants. This is possible because, while catalyzing mechanistically distinct reactions, the corresponding class II ent-copalyl diphosphate synthase and subsequently acting class I ent-kaurene synthase are clearly homologous. Nevertheless, the vast majority of the plant lower terpene synthases are significantly smaller than these diterpene synthases.

Structural characterization of plant and microbial lower terpene synthases demonstrated that the class I reaction is catalyzed in a conserved α-helical bundle domain,
that has been termed the α domain. This includes the placement of two highly conserved acidic motifs, DDXXD and (N,D)DXX(S,T)XXXE, which coordinate the trio of Mg\(^{2+}\) required for catalysis. However, the plant enzymes seem to invariably contain an additional N-terminal domain that appears to be relictual. In particular, this clearly is derived from the ancestral diterpene synthases, although these enzymes further contain an additional N-terminal large sequence element as well, with these domains termed β and γ, respectively.

The β domain in class II diterpene cyclases contains a DXDD motif required for such catalysis. This invokes similarities to the squalene-hopene cyclase (SHC) that also catalyzes such energetically difficult C=C protonation-initiated cyclization using an analogous motif. Furthermore, structural homology of this β domain, as well as the positioning and catalytic role of the DXDD motifs, has been noted between these seemingly unrelated but mechanistically similar enzymes. Conservation of the additional γ domain in class II diterpene cyclases has been used to argue that this also might be involved in class II catalysis, drawing further parallels to the triterpene cyclases, which are composed of a bidomain γβ structure wherein the active site is located between these domains. This then suggests that the γβα domain structure of eukaryotic diterpene synthases originated from functional fusion of separate, smaller class II (γβ domain) and I (α domain) enzymes, such as those found in bacteria. Consistent with this hypothesis is the presence of bifunctional diterpene synthases in fungi, as well as gymnosperms and lower plants, which further have been suggested to closely resemble the ancestral plant terpene synthase, along with the structure of taxadiene synthase from *Taxus brevifolia* (TbTS), although this is a strictly
class I diterpene synthase. One such bifunctional enzyme is abietadiene synthase from *Abies grandis* (AgAS), which is involved in grand fir tree resin acid biosynthesis (Figure 1). AgAS was the first bifunctional diterpene synthase to be cloned, and has been extensively characterized; including labeling and reaction intermediate analog binding studies, as well as mutational analysis of catalysis, and identification of single residue switches for both product outcome and biochemical regulation via susceptibility to Mg$^{2+}$-dependent inhibition. Here we report the crystal structure of AgAS at 2.3 Å resolution, which, along with that of the ent-copalyl diphosphate synthase from *Arabidopsis thaliana* (AtCPS) involved in gibberellin phytohormone biosynthesis reported in the accompanying paper, provide the first structures for class II diterpene cyclases. The resulting insights into terpene synthase evolution and details of the catalyzed reactions are discussed.

**Results**

The structure of AgAS is composed of three α-helical domains (α, β, and γ), much as previously predicted, and resembling that just reported for TbTS, as well as that reported in the accompanying paper for AtCPS. As a bifunctional diterpene synthase, AgAS differs from both TbTS and AtCPS in that both the class I and II active sites are functional, rather than the singular catalytic sites found in these monofunctional diterpene synthases. Accordingly, both are found as deep clefts defined by the presence of the relevant catalytic motifs. Notably, there is clear structural homology between the γβ domains of AgAS and the bidomain structure of the triterpene cyclases, as their structures can be superimposed on each other with a Cα backbone RMSD of 2.9 Å (Figure 2B). This supports our previously
advanced suggestion that class II diterpene cyclases evolved from the mechanistically similar triterpene cyclases.

The observed structure also is consistent with our previous suggestion that plant diterpene synthases originated from fusion of bacterial type class II and I diterpene synthases. This ancestral bifunctional diterpene synthase was presumably involved in gibberellin biosynthesis, and underwent gene duplication and subfunctionalization to give rise to the separate class II and class I enzymes that carry out such biosynthesis in higher plants. Given the ancestral role for diterpene synthases in the extensive plant terpene synthase family, the extended structural arrangement observed with AgAS provides some further insight into this evolutionary process. In particular, the β domain is sandwiched between the α and γ domains, such that these do not interact with each other, while the β domain forms extensive interactions with both (Figure 2A). Surprisingly, despite their cooperative formation of the class II active site, the interactions between the β and γ domains may have less structural importance, as suggested by the fact that the β domain is retained in all plant terpene synthases, while the γ domain appears to be relatively readily lost in monofunctional class I terpene synthases.

The class I active site is located in the α domain, which is clearly homologous to that of other class I terpene synthases (Figure 2C). Nevertheless, our structural characterization enables visualization of a previously described single residue switch for product outcome. In particular, the relevant residue, Ala723 in AgAS, which is found lining the active site underneath the acidic Mg\(^{2+}\) binding motifs. We have previously hypothesized that this side chain is in close proximity to the carbocation in the isopimar-15-en-8-yl\(^+\) intermediate.
formed by initial cyclization, such that substitution with a hydroxyl containing Ser stabilizes this sufficiently long enough for deprotonation\textsuperscript{47}, yielding the isopimara-7,15-diene observed as the major product of the AgAS:A723S mutant\textsuperscript{45}. Consistent with this, automated docking of sandaracopimaradiene (a minor product of both wild-type and AgAS:A723S) results in location of the relevant C8 near this residue (Figure 3).

The class II active site sits between the γ and β domains, as strongly suggested by our AgAS structure, and demonstrated by the location of the class II diterpene cyclase reaction analog 14,15-dihydro-15-azageranylgeranyl-S-thiolodiphosphate co-crystallized with AtCPS\textsuperscript{45}. This is analogous to the location of the triterpene cyclase active site, and such structural homology further extends to the identity and arrangement of catalytic residues. In particular, there is very similar positioning of the DXDD motifs in AgAS and SHC. In both the “middle” aspartic acid acts as the general acid\textsuperscript{24,48}, and in SHC this has been shown to be activated by interactions with a histidine\textsuperscript{26,48}. This histidine aligns with an asparagine that is conserved in all plant class II diterpene cyclases, and the corresponding N451 forms a hydrogen bond with the catalytic D404 in AgAS (Figure 4). Mutation of this asparagine to an alanine (AgAS:N451A) reduces catalytic activity >100-fold, consistent with a role in catalysis for this residue. Furthermore, both SHC and AgAS use interactions with a water molecule to impose an anti-orientation of the proton on the catalytic “middle” aspartic acid, which increases its acidity.

Two well-conserved aromatic residues line the AgAS class II active site near the catalytic aspartic acid (F354 and W358), and these seem likely to be involved in stabilizing carbocation reaction intermediates via interactions with their π quadrupoles, which face the active site. Indeed, it has already been demonstrated that W358 is important for catalysis, as
AgAS:W358A exhibits a 1,000-fold decrease in catalytic activity\(^\text{23}\). Furthermore, this tryptophan is completely conserved in all class II diterpene cyclases, while the neighboring F354 is similarly well-conserved (i.e., is always found as either as a phenylalanine or tyrosine), and these seem to correspond to a catalytically important tryptophan (W312 in SHC) in triterpene cyclases\(^\text{49}\).

Intriguingly, these two aromatic residues flank a residue that our previous results indicate controls the susceptibility of class II diterpene cyclases to Mg\(^{2+}\)-dependent inhibition\(^\text{44}\). While we originally hypothesized that this residue, conserved as a histidine or arginine in all plant class II diterpene cyclases, would be directly involved in catalysis, this clearly is not the case. Although mutation of the corresponding histidine in AtCPS to alanine reduces the catalytic rate >1,000-fold\(^\text{44}\), the corresponding R356 in AgAS is found on the opposite side of the relevant helix and reaches to the protein surface rather than forming any part of the active site (Figure 5). Nevertheless, mutation of this to alanine (AgAS:R356A) similarly reduces the catalytic rate 1,000-fold. Thus, we hypothesize that the observed dramatic effects from loss of this basic residue are exerted via changes in the positioning of the flanking aromatic residues.

Notably, this residue is specifically conserved as a histidine in all plant class II diterpene cyclases involved in gibberellin phytohormone biosynthesis, such as AtCPS, but is an arginine in all such enzymes dedicated to more specialized/secondary metabolism, such as AgAS. Furthermore, while AtCPS is susceptible to Mg\(^{2+}\)-dependent inhibition, AgAS is not\(^\text{50}\). Strikingly, interchanging the identity of this basic residue is sufficient to interchange the susceptibility of AtCPS and AgAS to Mg\(^{2+}\)-dependent inhibition\(^\text{44}\). The mechanism underlying
this remarkable conservation pattern and regulatory effect on catalysis is unclear, but seems likely to also involve the neighboring aromatic residues.

Unfortunately, the role of Mg$^{2+}$ in class II diterpene cyclases is not clarified by our structure, as no metal ions are present. While optimal activity requires the presence of Mg$^{2+}$, some can be observed even in the absence of any divalent metal ion$^{21,50}$, leaving its role unclear. We have previously hypothesized that Mg$^{2+}$ serves as a co-factor or co-substrate bound to the diphosphate moiety to help position GGPP for cyclization$^{21,50}$, and have speculated that this might involve a weakly conserved E$^{223}$DXXD motif$^{19}$. However, these residues are found splayed apart, and quite distant from the catalytic center defined by the DXDD motif (i.e., on or outside the rim of the class II active site cleft). Thus, it remains unclear how Mg$^{2+}$ exerts either activating or inhibitory effects, and further work will be necessary to elucidate the mechanisms underlying these.

**Discussion**

The structure of AgAS reported here provides insights into both the enzymatic mechanism and evolution of terpene synthases. In particular, the observed structural homology of the class II active site containing γβ domains with triterpene cyclases demonstrates common origins for these mechanistically similar enzymes. This extended to active site arrangement, which led to the identification of new catalytic residues (i.e., N451) and visual rationalization for others (i.e., W358). Furthermore, while exhibiting the expected homology for the class I active site containing α domain, the structure still enabled further analysis of the previously identified single residue switch for product outcome. In addition, the overall arrangement of the γβα domain structure provided a rationale for the observed βα
domain structure of most plant lower terpene synthases. Thus, coupled to the extensive previous studies of this model diterpene synthase, the AgAS structure not only provides insights into the class II and class I diterpene cyclization reactions catalyzed by this bifunctional enzyme, but the evolution of terpene synthases more generally.

Methods

Cloning, expression and purification of pseudomature abietadiene synthase.

The Δ84 pseudomature form of AgAS was recombinantly expressed using a previously described pSBET construct*, typically in the C41 OverExpress strain of *Escherichia coli* (Lucigen) grown in NZY media. For selenomethionine labeling, AgAS was expressed in B834(DE3) methionine auxotrophic *E. coli* (Novagen) grown in Selenomethionine Medium Base plus Nutrient Mix (Molecular Dimensions) with the addition of *L*-selenomethionine (Fisher). AgAS was then purified much as previously described *. Briefly, clarified bacterial lysates were initially fractionated over type II ceramic hydroxyapatite, and AgAS purified over Mono Q and type II ceramic hydroxyapatite again using a BioLogic LP system (Bio-Rad), before a final polishing purification step over a Mono Q column on an ÄKTAfplc system (Pharmacia). The resulting AgAS (>98% pure by SDS-PAGE) was then dialyzed against protein storage buffer (10 mM Bis-Tris, pH 6.8, 10% (v/v) glycerol, 150 mM KCl, 10 mM MgCl₂, and 5 mM DTT), following which it could be stored at -80 °C for several months without significant loss of activity.

Crystallization of AgAS.

Initial crystallization trials were carried out at the Hauptman-Woodward Institute. Promising conditions were replicated in our lab by the hanging-drop vapor diffusion methods at both 4 and 18 °C. Typically, a 2 µL drop of AgAS in storage buffer was mixed with 1.6
μL drop of precipitant solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), as well as with 0.4 μL additional 0.1 M L-proline or 0.1 M phenol from additive screen (Hampton Research), and equilibrated against 1 mL of precipitant solution in the well. Plate-shaped crystals with maximal dimensions of 0.3 x 0.3 x 0.05 mm appeared within a week. Crystals were briefly transferred to a cryoprotectant solution (25% PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate, 10% (v/v) glycerol), and then flash-frozen.

**X-ray diffraction data collection, processing, phasing and model refinement.**

X-ray diffraction data were collected at the Advanced Photon Source, that for crystal I on beamline 24-ID-D, and that for crystal II on beamline 19-ID, using λ = 0.9792 Å, corresponding to the absorption peak of selenium. All diffraction data were indexed and merged using HKL-3000. Selenium sites in crystal I (located by the Solve routine in Phenix), led to the automatic generation of several peptide fragments by Autosol. A complete model (R_free ~ 0.29, resolution of 2.8 Å) for crystal I resulted from cycles of model building (Xtalview) and restrained refinement by Refmac and CNS. The model from crystal I enabled a molecular replacement structure determination of crystal II using Phaser. CNS refinement of the structure from crystal II was refined against data to 2.3 Å resolution using standard restraints on stereochemistry and thermal parameters, and auxiliary restraints on donor-acceptor backbone distances in regions of regular secondary structure. Statistics for the refined crystal II structure are in Table 1, with Ramachandran plot statistics from PROCHECK. The four, monomers in the asymmetric unit are nearly identical, exhibiting pairwise RMSD superpositions of Cα atoms of 0.25 Å. For each monomer, electron density for residues of the N-terminus (M84-A109) is absent, although sequencing (Edman
degradation performed by the Iowa State University Protein Facility) indicated a mixture of polypeptides beginning with residues 84 or 85. Electron density for residues 846-849 is absent in chains A and D, but present, albeit weakly, in chains B and C. All monomers contained density in the class II active site cleft, which was fit as tri-peptides. All protein structure figures were generated with the B chain model of AgAS, using PyMOL (www.pymol.org), with product docking (where applicable) carried out on a model without the tri-peptide in the class II active site at www.dockingserver.com.

**Cloning, expression and purification of His tagged mutants of rAgAS**

AgAS mutants (R356A and N451A) were constructed via overlapping PCR, then expressed as 6×His tagged fusion proteins, purified, and kinetically characterized, just as previously described ⁴⁴.

**Acknowledgements**

This work was supported by a grant from the NIH to R.J.P. (GM076324). Additionally, we thank Prof. Gaya Amarasinghe (Iowa State University) for collecting diffraction data on crystal I, and the Argonne National Laboratory for beamline access, which was supported by DOE (W-31-109-ENG-38 and DE-AC02-06CH11357).

**References**


Peters, R.J. & Croteau, R.B. Abietadiene synthase catalysis: Conserved residues


39 Peters, R.J. et al. Abietadiene synthase from grand fir (*Abies grandis*):


Figures

Figure 1: Separate class II and class I cyclization reactions catalyzed by AgAS, including key isopimar-15-en-8-yl class I reaction intermediate.
Figure 2: AgAS crystal structure and comparison to known terpene synthases. (A) Cartoon diagram of AgAS structure with γβα domains labeled. (B) Superposition of AgAS γβ domains (green) on SHC (blue). (C) Superposition of the AgAS α domain (green) on that of the sesquiterpene cyclase 5-epi-aristolochene synthase (blue).
**Figure 3:** AgAS class I active site, with docked sandaracopimaradiene and pyrophosphate coproduct, as well as trio of magnesium ions. Also shown is the Ala side chain corresponding to the single residue switch for product outcome, to demonstrate its proximity to C8 of the sandaracopimaradiene, which corresponds to the location of the carbocation in the relevant isopimar-15-en-8-yl$^+$ reaction intermediate.
**Figure 4:** Catalytic residues from the AgAS class II active site (green), superimposed on their equivalents in SHC (light blue), with those from AgAS labeled.
**Figure 5:** Regulatory switch residue (Arg356 in AgAS) lies outside of class II active site (defined by catalytic DXDD motif and interacting Asn).
Table 1

<table>
<thead>
<tr>
<th>Table 1. Data collection and refinement statistics</th>
<th>rAgAS (crystal I)</th>
<th>rAgAS (crystal II)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Se Peak)</td>
<td>0.9793</td>
<td>0.9792</td>
</tr>
<tr>
<td>Resolution Range (Å)</td>
<td>50.00 - 2.50</td>
<td>50.00 - 2.30</td>
</tr>
<tr>
<td>No. of reflections (total/unique)</td>
<td>271627/61432</td>
<td>474752/148465</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.2 (98.4)</td>
<td>97.7 (96.6)</td>
</tr>
<tr>
<td>I/σ²</td>
<td>25.1 (2.4)</td>
<td>18.5 (3.3)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.104 (0.739)</td>
<td>0.092 (0.360)</td>
</tr>
<tr>
<td>FOM for SAD</td>
<td>0.46</td>
<td>---</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rwork/Rfree c</td>
<td>0.208 / 0.287</td>
<td>0.196 / 0.250</td>
</tr>
<tr>
<td>Protein atoms d</td>
<td>---</td>
<td>24815</td>
</tr>
<tr>
<td>Solvent atoms d</td>
<td>---</td>
<td>1346</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.008</td>
<td>0.006</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>1.19</td>
<td>1.10</td>
</tr>
<tr>
<td>Average B factors (Å²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein chains</td>
<td>---</td>
<td>34.3</td>
</tr>
<tr>
<td>Solvent</td>
<td>---</td>
<td>35.3</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>---</td>
<td>91.5</td>
</tr>
<tr>
<td>Additionally allowed (%)</td>
<td>---</td>
<td>8.4</td>
</tr>
<tr>
<td>Generously allowed (%)</td>
<td>---</td>
<td>0.1</td>
</tr>
<tr>
<td>Disallowed (%)</td>
<td>---</td>
<td>0</td>
</tr>
</tbody>
</table>


- Number in parentheses refer to the outer 0.04 Å shell of data.
- $R_{merge} = \sum |I - \langle I \rangle| / \sum I$, where $I$ is the observed intensity and $\langle I \rangle$ is the average intensity calculated from replicate data.
- $R_{work} = \sum |F_o - F_c| / \sum |F_o|$ for reflections contained in the working set, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively, and $R_{free}$ is calculated in the same way for reflections contained in the test set held aside during refinement.
- Per asymmetric unit.
Chapter III: Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants

This paper was published in and reprinted from phytochemistry, 2009 Feb;70(3):366-9.

Ke Zhou, Reuben J. Peters *

Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Molecular Biology Building, Rm. 4216, Ames, IA 50011, United States

Abstract

Terpene synthases (TPS) require divalent metal ion co-factors, typically magnesium, that are bound by a canonical DDXXD motif, as well as a putative second, seemingly less well conserved and understood (N/ D)DXX(S/T)XXXE motif. Given the role of the Ser/Thr side chain hydroxyl group in ligating one of the three catalytically requisite divalent metal ions and the loss of catalytic activity upon substitution with Ala, it is surprising that Gly is frequently found in this ‘middle’ position of the putative second divalent metal binding motif in plant TPS. Herein we report mutational investigation of this discrepancy in a model plant diterpene cyclase, abietadiene synthase from Abies grandis (AgAS). Substitution of the corresponding Thr in AgAS with Ser or Gly decreased catalytic activity much less than substitution with Ala. We speculate that the ability of Gly to partially restore activity relative to Ala substitution for Ser/Thr stems from the associated reduction in steric volume enabling a water molecule to substitute for the hydroxyl group from Ser/Thr, potentially in a divalent metal ion coordination sphere. In any case, our results are consistent with the observed conservation pattern for this putative second divalent metal ion binding motif in plant TPS.

1. Introduction

Terpene synthases (TPS) catalyze complex cyclization and/or rearrangement of isoprenyl
pyrophosphate precursors (Christianson, 2006). This creates the diverse hydrocarbon skeletal/backbone structures that underlie the amazing variety of terpenoid natural products, and often comprises the committed step in particular biosynthetic pathways. TPS typically ionize the allylic pyrophosphate ester linkage of their substrate to initiate carbocationic reactions (i.e. class I TPS; EC 4.2.3.-). Catalysis is dependent on divalent metal ion cofactors, usually Mg$^{2+}$, which were hypothesized to bind to and assist ionization of the pyrophosphate. Sequence analysis of initially cloned class I TPS led to the additional hypothesis that a DDXX(D/E) motif was involved in divalent metal binding based on previous findings in the mechanistically similar prenyltransferases (Facchini and Chappell, 1992). These hypotheses have been confirmed by mutational and structural analysis, and the DDXX(D/E) motif is now considered a characteristic feature of class I TPS (Christianson, 2006).

In addition to this DDXX(D/E) motif, based on the structure of trichodiene synthase from the fungus Fusarium sporotrichioides, it has been hypothesized that class I TPS contain a second conserved metal binding motif, (N/D)DXX(S/T)XXXE (metal binding residues in boldface) (Rynkiewicz et al., 2001). These residues are generally conserved in class I TPS, and their importance for catalysis in plant class I TPS has been demonstrated by alanine scanning mutagenesis in the diterpene cyclase abietadiene synthase from Abies grandis (AgAS), wherein Ala substitution for any one of the identified metal binding residues reduces catalytic efficiency ~10,000–fold (Peters and Croteau, 2002). However, the ‘middle’ Ser/Thr position is not completely conserved, with Gly found in this position in a surprising number of plant class I TPS. For example, within the diterpene synthases that are a focus of our research, while AgAS contains Thr, the kaurene synthases from Cucurbita maxima (CmKS)
and Arabidopsis thaliana (AtKS) contain Ser and Gly, respectively, at the corresponding position (Fig. 1). Yet all three enzymes are catalytically active (Stofer Vogel et al., 1996; Yamaguchi et al., 1996; Yamaguchi et al., 1998). In the rice (Oryza sativa) family of eight functional kaurene synthase-like (OsKSL) class I diterpene cyclases, two have Gly in the otherwise conserved Thr/Ser position (Xu et al., 2007). Gly also is found in the Thr/Ser position in a number of plant mono-and sesqui-terpene class I TPS as well. Accordingly, investigators working with plant TPS often do not acknowledge this second motif. Nevertheless, the corresponding residues interact with a Mg$^{2+}$ ion in all of the plant class I TPS structures known to date (Hyatt et al., 2007; Kampranis et al., 2007; Starks et al., 1997; Whittington et al., 2002). Although in each case, this second motif is conserved (i.e. contains a Ser or Thr at the appropriate position).

AgAS is a bifunctional enzyme that catalyzes two sequential cyclization reactions in distinct active sites (Peters et al., 2001). This includes a canonical class I TPS domain with both divalent metal binding motifs, whose importance for catalysis has been previously demonstrated by alanine scanning mutagenesis (Peters and Croteau, 2002). To investigate the ability of Gly to functionally substitute for Ser/Thr in the putative second divalent metal binding motif, we replaced the Thr at this position in AgAS with Ser, Ala, or Gly. Kinetic assays specifically examining class I TPS activity revealed that substitution by Ser or Gly decreases the activity of AgAS much less than substitution with Ala. This observation is consistent with the observed conservation pattern in plant class I TPS. We further speculate that the ability of Gly to partially restore activity relative to Ala substitution for Ser/Thr stems from the associated reduction in steric volume enabling a water molecule to substitute for the hydroxyl group from Ser/Thr, potentially in a divalent metal ion coordination sphere.
2. Results

2.1. Development of a coupled assay to examine AgAS class I activity

AgAS catalyzes protonation-initiated cyclization (i.e. a class II TPS reaction) of the universal diterpene precursor GGPP (1) to CPP (2) prior to catalyzing class I cyclization of CPP (2) to abietadienes in a separate active site (Fig. 2). To selectively assay class I activity, CPP (2) was produced from GGPP (1) with a mutant AgAS wherein the first Asp of the DDXXD motif has been substituted by Ala (AgAS:D621A), which essentially completely eliminates class I activity (~10⁶ reduction) (Peters et al., 2001). The resulting CPP (2) was then utilized to selectively assay AgAS class I activity. These assays were carried out using recombinant AgAS encoded with an N-terminal 6 x His tag for ease of purification. Because it has been shown that a KR motif at the N-terminus of AgAS forms part of the class I active site (Peters et al., 2003), the 6 x His tag was separated from the beginning of the AgAS protein sequence by a 25 amino acid linker peptide. This change in protein structure did lead to a reduction in catalytic efficiency, with an increase in KM from 0.4 to 1.2 uM and decrease in kcat from ~ 2 to 0.4 s⁻¹, resulting in an overall ~17-fold decrease in kcat/KM from 5 to 0.3(X10⁶)M⁻¹s⁻¹ relative to untagged AgAS. Nevertheless, given the ease of purification and relative nature of mutant analysis, we chose to carry out our studies in the context of this 6 x His-tag.

2.2. Mutational analysis of the Ser/Thr position in the second TPS divalent metal binding motif

Given the observed conservation pattern and previous mutational analysis of the Ser/Thr position in the (N/D)DXX(S/T)XXXE second divalent metal binding motif, Ser, Ala, or Gly
were substituted for the corresponding Thr769 in recombinantly 6 x His tagged AgAS, creating AgAS:T769S, AgAS:T769A, and AgAS:T769G, respectively. Much as previously reported (Peters and Croteau, 2002), AgAS:T769A exhibits a large reduction in catalytic activity, with kcat reduced to $< 10^{-4}$ s$^{-1}$. By contrast, AgAS:T769S and AgAS:T769G retain more significant amounts of catalytic activity (Table 1). Ser substitution reduces kcat ~13-fold, from 0.4 to 0.03 s$^{-1}$, with little effect on the pseudo-substrate binding constant $K_M$, for only a ~17-fold reduction in catalytic efficiency. Even Gly substitution has relatively little effect, reducing kcat ~ 130-fold, to 0.003 s$^{-1}$, again with little effect on $K_M$, for an overall ~167-fold reduction in catalytic efficiency. Again as previously reported (Peters and Croteau, 2002), there is a slight change in product profile with the AgAS:T769A mutant, which exhibits some redistribution in the ratio of various abietadiene double bond isomers. However, substitution with Ser or Gly did not lead to significant changes in product profile (data not shown). Finally, comparison of the circular dichroism spectra for wild type and mutant AgAS indicate that these substitutions do not significantly alter protein structure (Fig. 3).

3. Discussion

Previous mutational analysis of the divalent metal binding motifs in AgAS (Peters and Croteau, 2002), as well as an aristolochene (sesquiterpene) synthase from the fungus Penicillium roqueforti (Felicetti and Cane, 2004), only examined the effect of Ala substitution for the Ser/Thr position in the (N/D)DXX(S/T)XXXE putative second TPS divalent metal binding motif. More recently, a detailed structure-function investigation of this motif in the trichodiene (sesquiterpene) synthase from F. sporotrichioides reported that
Thr substitution for the Ser found at the Ser/Thr position leads to a ~700-fold decrease in catalytic efficiency. This largely stemmed from a relatively large 77-fold increase in KM, which crystallographic structure analysis suggested arises from steric incompatibility of the introduced c-methyl of the Thr side chain with other elements of the active site (note that Ser is typically found at this position in fungal class I TPS such as this) (Vedula et al., 2008). However, in contrast to these mutational analyses, Gly is occasionally found in the Ser/Thr position in functional class I TPS from plants (e.g. Fig. 1). Here we tried to reconcile these disparate findings by specifically examining the functional plasticity of the Ser/Thr position in the putative second TPS divalent metal binding motif of AgAS, a model plant diterpene synthase.

Our results demonstrate that, at least in the context of the AgAS active site, substitution of the Thr divalent metal ligand with Ser has relatively little effect on catalysis, either efficiency or reaction outcome. Selective removal of the directly ligating hydroxyl group by Ala substitution drastically reduces catalytic efficiency. However, complete side chain removal by Gly substitution partially restores catalytic activity. From these results, we speculate that replacement of Thr with Gly, but not Ala, creates sufficient space within the active site for a water molecule to bind and substitute for the Ser/Thr hydroxyl group, potentially as a divalent metal ion ligand (Fig. 4), albeit with significantly reduced efficiency. In particular, the methyl group of the Ala side chain would prevent a water molecule from occupying a position approximating that of a Ser/Thr hydroxyl group (i.e. the space occupied by the oxygen of such a hydroxyl group containing side chain are occluded by the methyl protons of Ala). We also note that the postulated water molecule would be shielded from the reactive carbocation intermediates by the divalent metal and ionized pyrophosphate complex,
consistent with the observed lack of hydroxylated products.

4. Concluding remarks

It has been noted that there are subtle, but functionally important, differences between the active sites of microbial and plant derived class I TPS (Christianson, 2006). Thus, given the absence of detailed structural information for AgAS, it is difficult to resolve the structure-function relationships underlying the differences between the results reported here and those reported for similar mutational analysis of the fungal trichodiene synthase (Vedula et al., 2008). While other deviations from this putative second TPS divalent metal binding motif can be found in plant TPS [e.g. see Martin et al. (2004), wherein several functional class I TPS contain a Gly in place of the ‘first’ N/D position], previous results (Peters and Croteau, 2002), along with those reported here, are consistent with an important role for this motif in plant, as well as fungal, TPS, despite a more divergent composition. Given the significant reduction in activity observed upon Gly substitution (Table 1), secondary changes are presumably required to restore full enzymatic activity. Nevertheless, our results provide biochemical observations consistent with the observed conservation pattern of this putative second TPS divalent metal binding motif in plants.

5. Experimental

5.1. General procedures

Unlabeled [E,E,E]-geranylgeranyl pyrophosphate (GGPP (1)) was purchased from Isoprenoids, LC (Tampa, FL). [1-3H]GGPP (1) was purchased from American Radiolabeled Chemicals (St. Louis, MO), Ni-NTA His-bind matrix was purchased from Novagen (Madison, WI), and molecular biology reagents were purchased from Invitrogen (Carlsbad, CA). Unless otherwise noted, all other chemicals were purchased from Fisher Scientific
5.2. Recombinant constructs

The recombinant pseudo-mature AgAS gene has been previously described (Peters et al., 2000), and transferred into the Gateway (Invitrogen) vector system via directional topoisomerization into pENTR/SD/D-TOPO (Cyr et al., 2007). Site-directed mutagenesis was carried out via PCR amplification of the pENTR/AgAS construct with overlapping mutagenic primers, and the mutant genes verified by complete sequencing. The resulting wild type and mutant genes were then transferred via directional recombination to the T7-promoter and N-terminal 6 x His fusion expression vector pDEST17. Use of the pDEST17 vector results in a 25 amino acid residue linker between the 6 x His tag and the cloned protein (here pseudo-mature AgAS). The linker peptide sequence is LESTSLYKKAGSAALFNFKKEPFT.

5.3. Enzyme expression and purification

Transformed bacterial cells including pDEST17/AgAS (wild type or one of the various mutants) were grown in 1 L of NZY medium at 37 °C with shaking to an A600 of 0.6–0.8, transferred to 16 °C for 1 h, then induced with 0.5 mM IPTG and incubated with shaking for an additional 16–20 h. Bacterial cells were harvested by centrifugation and resuspended in 20 mL lysis buffer (50 mM Bis–Tris, pH 6.8, 150 mM KCl, 10 mM MgCl2, 10% Glycerol). After sonication (Branson Sonifier 450: 3 x 5 s continuous output at a setting of 5), the lysate was clarified by centrifugation at 15,000g for 25 min at 4 °C. AgAS was purified using Ni-NTA His-bind resin in batch mode. In brief, clarified lysate was added to 1 mL of Ni-NTA that had been pre-equilibrated in wash buffer (50 mM Bis–Tris, pH 6.8, 1 mM DTT), 6 x His
tagged AgAS bound by gently shaking at 4 °C for 1 h, then washed with 50 mL wash buffer with 20 mM imidazole, and AgAS eluted with 2 x 1 ml elution buffer (50 mM Bis-Tris, pH 6.8, 250 mM imidazole, 1 mM DTT). The imidazole was removed by concentration through a 50 kDa molecular weight cutoff centricon (Millipore, Billerica, MA) to ~ 0.2 mL, followed by dilution in wash buffer to 2 mL. AgAS protein concentration was determined by A280 using the calculated extinction coefficient 138,350 M⁻¹ cm⁻¹. Structural integrity was measured by circular dichroism. For this purpose, AgAS was purified and analyzed at 0.5 mg/mL in 10 mM PBS, 10 mM MgCl₂, 1 mM DTT, 10% glycerol.

5.4. Enzymatic analysis

Kinetic analysis was carried out using the D621A mutant of AgAS, which eliminates class I activity (Peters et al., 2001), to convert [1-⁻³H]GGPP (1) to [1-⁻³H]copalyl pyrophosphate (CPP). In brief, 1 ml reactions in assay buffer (50 mM Hepes, pH 7.2, 0.1 mM MgCl₂, 5% glycerol, 5 mM DTT, 100 mM KCl and 0.1 mg/ ml a-casein) containing 50 lM [1-⁻³H]GGPP (1) and 1 lM AgAS:D621A were run for 2 h at 30 °C, which is sufficient to completely convert this amount of unlabeled GGPP to CPP. Kinetic assays were then performed similar to the method described previously (Peters et al., 2000). In brief, duplicate 1 ml reactions in assay buffer containing 3 nM wild type AgAS were initiated by the addition of [1-⁻³H]CPP (2), allowed to react for 1 min at 30 °C, then stopped by the addition of KOH to 0.2 M and EDTA to 15 mM. For analysis of the mutants, the enzymatic concentration was increased to 10, 50, or 100 nM, and incubation times increased to 5 min., 2 h, or 20 h for T769S, T769G, or T769A, respectively. The produced diterpenes were then extracted by hexane, the pooled extract passed over a short silica gel column, and product formation
assessed by scintillation counting. The resulting data was analyzed using Kaleidagraph (Synergy, Reading, PA).

Diterpene products were analyzed by co-expressing AgAS (wild type or one of the various mutants) with GGPP synthase in Escherichia coli, as previously described (Cyr et al., 2007). The resulting organic solvent extract of these recombinant cells were analyzed by gas chromatography with mass spectral detection (GC–MS), which was performed with a VF-1 column on a Varian (Palo Alto, CA) 3900 GC with Saturn 2100 ion trap MS system in electron ionization (70 eV) mode. Samples (1 uL) were injected in splitless mode at 50 °C and, after holding for 3 min. at 50 °C, the oven temperature was raised at a rate of 14 °C/min. to 300 °C, where it was held for an additional 3 min. MS data from 90 to 600 m/z were collected starting 12 min. after injection until the end of the run.

Acknowledgements

We thank Meimei Xu for assistance with molecular biology and Francis Mann for assistance with the GC–MS analyses. This work was supported by NIH grant GM076324 to R.J.P.

References


Figures

Fig. 1. Alignment of the TPS divalent metal binding motifs for selected diterpene synthases (above the alignment indicates the divalent metal binding residues; below, the ‘middle’ Ser/Thr position of the second binding motif investigated here).

|      | 621 | DDLVD |    | 765 | NDTKTVQAEP |      | 531 | DDPFD |    | 675 | NDIGFKRE |      | 536 | DDFYD |    | 680 | NDIRSYDRE |      |
|------|-----|-------|----|-----|------------|------|-----|-------|----|-----|----------|------|-----|-------|----|-----|------------|
| AgAS |     |       |    |     |            |      | AtKS|       |    |     |          |      | CmKS|       |    |     |            |      |
| OsKSL10|    |       |    |     |            |      | OsKSL11|      |    |     |          |      | OsKSL8|      |    |     |            |      |
| OsKSL15|    |       |    |     |            |      | OsKSL6|      |    |     |          |      | OsKSL7|      |    |     |            |      |
| OsKSL4|    |       |    |     |            |      |      |       |    |     |          |      |      |       |    |     |            |      |

*
Fig. 2. Sequential cyclization reactions catalyzed by AgAS.
Fig. 3. Circular dichroism spectra for wild type AgAS (solid line) and AgAS:T769G (dotted line).
Fig. 4. Hypothesized interaction of various side chains in the ‘middle’ position of the second TPS divalent metal binding site.

Thr  Ser  Ala  Gly + H₂O
Table 1

Kinetic parameters.

<table>
<thead>
<tr>
<th>Abietadiene synthase</th>
<th>kcat (s⁻¹)</th>
<th>KM (μM)</th>
<th>kcat/KM (⁻¹05 M⁻¹ s⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>0.4 ± 0.1</td>
<td>1.2 ± 0.3</td>
<td>3</td>
</tr>
<tr>
<td>T769S</td>
<td>0.03 ± 0.01</td>
<td>1.5 ± 0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>T769A</td>
<td>&lt;10⁻⁴</td>
<td>NDa</td>
<td>&lt;0.0001a</td>
</tr>
<tr>
<td>T769G</td>
<td>0.003 ± 0.001</td>
<td>1.5 ± 0.4</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*Previous analysis indicates T769A increases KM 5-fold (ND, not determined).*
Chapter IV: Insights into the evolution of diterpenoid metabolism in cereal crop plants from functional characterization of wheat diterpene synthases

Munuscript in preparation and to be submitted to Phytochemistry

Ke Zhou¹, Meimei Xu¹, Mollie Tiernan¹, Qian Xie¹, Tomonobu Toyomasu², Chizu Sugawara², Madoka Oku², Masami Usui², Wataru Mitsuhashi², Peter M. Chandler³, Reuben J. Peters¹,*

¹Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, IA 50011, U.S.A.; ²Department of Bioresource Engineering, Yamagata University, Tsuruoka, Yamagata, 997-8555, Japan; ³CSIRO Plant Industry, GPO Box 1600 Canberra, ACT 2601, Australia

Abstract

Wheat (Triticum aestivum) and rice (Oryza sativa) are two of the most agriculturally important cereal crop plants. Rice is known to produce numerous diterpenoid natural products that serve as phytoalexins and/or allelochemicals. Specifically, these are labdane-related diterpenoids, derived from a characteristic labdadienyl/copalyl diphosphate (CPP), and whose biosynthetic relationship to gibberellin biosynthesis is evident from the relevant expanded and functionally diverse family of ent-kaurene synthase-like (KSL) genes found in rice (OsKSL). Here we report biochemical characterization of a similarly expansive family of KSL from wheat (the TaKSL). In particular, beyond ent-kaurene synthases (KS), wheat also contains several functionally diversified KSL. These react either with the ent-CPP intermediate common to gibberellin biosynthesis or with the normal stereoisomer of CPP that also is found in wheat (as demonstrated by the accompanying description of wheat CPP synthases). Comparison with a barley (Hordeum vulgare) KS indicates conservation of monocot KS, with early and continued expansion and functional diversification of KSL in at
least the small grain cereals. In addition, the ability of some of the normal CPP reactive TaKSL to also react with syn-CPP echoes previous findings with the OsKSL family, with such enzymatic promiscuity/plasticity providing insight into the evolution of diterpenoid metabolism, at least in the cereal crop plant family, which is discussed here.

1. Introduction

Cereal crop plants provide the bulk of the world's caloric intake, with wheat and rice representing the two most important for direct human consumption. Rice has served as a model for the cereal crop plant family, as its agricultural importance and relatively small genome size led to early and thorough sequencing (Goff et al., 2002; Yu et al., 2002), which was complimented by a large scale cDNA sequencing effort (Kikuchi et al., 2003). The resulting complete gene list has enabled comprehensive investigation of various aspect of rice physiology, including metabolism. Such work then provides the basis for similar investigations in other cereal crops such as wheat.

Rice is a particularly prolific producer of labdane-related diterpenoids, which have been suggested to as phytoalexins in defense against microbial pathogens and as allelochemicals suppressing the growth of neighboring weed plants (Peters, 2006; Toyomasu, 2008). The biosynthesis of this super-family of natural products is derived from that of the gibberellin phytohormones, at least in plants, and characterized by a similar pair of sequentially catalyzed cyclization reactions (Peters, 2010). In particular, bicyclization of the general diterpenoid precursor \((E,E,E)\)-geranylgeranyl diphosphate (GGPP), typically to a labdadienyl/copalyl diphosphate (CPP) intermediate, which is then often further cyclized to a hydrocarbon olefin. The corresponding enzymes have been termed CPP synthases (CPS) and
ent-kaurene synthase-like (KSL), respectively, for their relationship to those found in all plants for gibberellin biosynthesis (Peters, 2006).

The rice CPS (OsCPS) and KSL (OsKSL) have been extensively investigated, with biochemical function assigned to each (Cho et al., 2004; Kanno et al., 2006; Nemoto et al., 2004; Otomo et al., 2004a; Otomo et al., 2004b; Prisic et al., 2004; Wilderman et al., 2004; Xu et al., 2004; Xu et al., 2007). In addition, consistent with a role in phytoalexin biosynthesis, many of the OsKSL exhibit inducible gene transcription (Peters, 2006; Toyomasu, 2008). Previous work with maize (Zea mays) demonstrated the inducible production of labdane-related diterpenes in this cereal crop plant (Mellon and West, 1979), and it has been suggested that CPS gene expansion and functional diversion to secondary/more specialized metabolism occurred early in the cereal crop plant family – i.e., the Poaceae (Prisic et al., 2004). Thus, it seems likely that more specialized labdane-related diterpenoid metabolism will be widespread throughout the Poaceae. Indeed, gene probing/mapping experiments using a KSL from barley (Hordeum vulgare) suggests that at least barley and wheat contain expanded CPS and KSL gene families (Spielmeyer et al., 2004). Some preliminary analysis of wheat CPS (TaCPS) has been previously reported (Toyomasu et al., 2009). However, only the ent-kaurene synthase (KS) activity expected for the requisite gibberellin biosynthesis has been previously reported from wheat (Aach et al., 1995). Here we report cloning and functional characterization of seven members of the wheat KSL family (TaKSL), along with demonstrating KS activity for the previously isolated barley KSL (HvKS), and discuss the implications of our findings for the evolution of diterpenoid metabolism in the cereal crop family, as well as more generally.

2. Results
2.1 Identification of kaurene synthase-like genes in wheat and barley

Isolation of HvKS has been previously reported (Spielmeyer et al., 2004). TaKSL were initially identified by homology searches against the available EST data with the known OsKSL. The corresponding full-length cDNA were then cloned by RT-PCR and, where necessary RACE. Through this effort, five clearly full-length KSL were found; TaKSL1, 2, 3, 4, and 6, which encode proteins of 837, 853, 856, 837, and 852 amino acid (aa) residues, respectively. In addition, two closely related genes (94% identical at the nucleotide sequence level) were cloned and assigned as TaKSL5-1 and TaKSL5-2 under the assumption that these are homoeologs (e.g., the most notable difference is an in-frame 63 nucleotide deletion in TaKSL5-2 relative to TaKSL5-1). Notably, as previously reported for TaKSL5-2 (Hillwig et al., submitted), TaKSL5-1 and TaKSL5-2 are substantially shorter than other KSL, resembling plant mono- and sesqui- terpene synthases in length. Specifically, these encode proteins of 663 and 641 aa in length, respectively, although they otherwise closely align with the other TaKSL (Fig. 1).

2.2 Induction of transcription of wheat labdane-related diterpene synthase genes

In rice, mRNA levels of the OsKSL, as well as OsCPS, involved in phytoalexin biosynthesis were dramatically increased by UV-irradiation (Peters, 2006). Thus, the possibility that mRNA levels of some of the TaKSL would be induced in response to UV-irradiation was analyzed by qRT-PCR. Indeed, mRNA levels of TaKSL1, TaKSL2, and TaKSL5, were found to be higher in UV-irradiated leaves, although those of TaKSL3, TaKSL4, and TaKSL6 were not (Fig. 2).

2.3 Biochemical characterization
The TaKSL and HvKS were functionally characterized via use of a previously developed metabolic engineering system that enables co-expression of both GGPP synthase and CPS producing the three commonly found stereoisomers of CPP, along with downstream KSL such as these (Cyr et al., 2007). This enabled analysis of their activity with not only the clearly physiologically relevant ent- (2) and normal (3) CPP (Wu et al., submitted), but that of syn- (4) stereochemistry as well (Fig. 3). Using this approach, HvKS was demonstrated to selectively react with ent-CPP (2) and produce only ent-kaurene (5). Similarly, the closely related TaKSL6 specifically reacts with 2 to produce 5. In addition, the putative homoeologs TaKSL5-1 and TaKSL5-2 also selectively react with 2 to chiefly produce 5, although both also produce ent-beyerene (6), in a ~3:1 ratio. As previously reported for TaKSL5-2 (Hillwig et al., submitted), TaKSL5-1 and TaKSL5-2 are substantially shorter than other KSL, and the most active constructs were those which resemble plant mono- and sesqui- terpene synthases. TaKSL4 reacts with normal CPP (3) to produce pimara-8(9),15-diene (7), and also will react with syn-CPP (4) to produce a number of diterpenes, chiefly syn-pimara-9(11),15-diene (8), along with at least five other relatively minor unidentified products. TaKSL3 exhibits relatively low activity, although does reacts exclusively with ent-CPP (2) to generate two diterpene products. Unfortunately, due to the low yield of those products, it was not possible to obtain enough of these compounds for structural characterization. TaKSL2 similarly exhibits low productivity, although it was possible to demonstrate that this enzyme will not only react with ent-CPP (2) to produce the known ent-pimara-8(14),15-diene (9), but also can utilize normal CPP (3) to make abietadiene (10). Interestingly, this activity was only detected with an extensively truncated construct resembling those for TaKSL5-1 and TaKSL5-2 (i.e., also similar to plant mono- and sesqui- terpene synthases). Finally, TaKSL1 reacts with
normal CPP (3) to produce iso-pimara-7,15-diene (11), and also will react with syn-CPP (4) to produce syn-iso-pimara-7,15-diene (12).

2.4 Molecular phylogenetic analysis of cereal KS(L)

Previous molecular phylogenetic analysis of cereal CPS indicates that expansion and functional divergence of at least two copies of this enzymatic gene to more specialized metabolism occurred prior to the speciation event separating the wheat and rice lineages (Toyomasu et al., 2009). To determine if similar early multiplication and functional divergence occurred with the KSL, as well as provide insight into potential physiological function of the various TaKSL family members, we carried out molecular phylogenetic analysis of the cereal KS(L). Specifically, we aligned the full-length amino acid sequences of all the OsKSL with HvKS and all the TaKSL but TaKSL5-1 and TaKSL5-2, which were excluded on the basis of their significant difference in size, along with the KS from the dicot Arabidopsis thaliana (AtKS). AtKS was then designated the outgroup sequence for the resulting phylogenetic tree, which was constructed using the nearest neighbor joining method (Fig. 4).

While the TaKSL and OsKSL largely cluster independently, indicating continued expansion and divergence of this gene family after separation of the wheat/rice lineages, there is some overlap. In particular, OsKSL4 clusters with TaKSL1 and TaKSL4, suggesting early KSL gene duplication and functional divergence in the small grain cereal lineage. Given the inducible transcription of at least TaKSL1 and OsKSL4, it seems likely that the ancestral enzymatic gene was similarly involved in more specialized metabolism. In addition, consistent with conservation of KS activity, HvKS and TaKSL6 are the most closely related KSL (sharing 91% aa sequence identity), and both are quite similar to OsKS1 (~68% aa
identity), which is required for gibberellin biosynthesis in rice (Sakamoto et al., 2004). This conservation pattern reflects the underlying evolutionary separation of rice, which falls into the Poaceae subfamily Oryzoideae, from wheat and barley, which fall within the separate Pooidae subfamily (Kellogg, 1998). Thus, it seems likely that HvKS and TaKSL6 are similarly involved in gibberellin biosynthesis, although this remains to be demonstrated. Consistent with this hypothesis, the highly divergent sequences, including overall length, as well as reduced product fidelity and inducible gene transcription, of TaKSL5-1 and TaKSL5-2 suggests that these are not involved in gibberellin metabolism, and these have been shown elsewhere to cluster with the rice pseudogenes OsKSL2 and OsKSL3, rather than OsKS1 (Hillwig et al., submitted).

Notably, while the rice KSL have clearly undergone repeated evolutionary derivation of substrate specificity (e.g., the syn-CPP specific OsKSL4 and OsKSL11 fall into separate clusters), the TaKSL from in this study do exhibit such functional conservation. In particular, the ent-CPP specific TaKSL cluster together, with the normal/syn-CPP specific TaKSL1 and TaKSL4 falling into a separate cluster. This latter cluster further includes the OsKSL4 that exhibits a similar substrate range (Morrone et al., in press), suggesting an early origin for such divergent substrate stereo-specificity, which is consistent with early diversification of the responsible CPS suggested in the accompanying report (Wu et al., submitted). However, there is no clear phylogenetic relationship among the inducible versus non-inducible TaKSL, which also was noted for the rice KSL (Peters, 2006). For example, while TaKSL1 and TaKSL2 exhibit UV-inducible transcription, TaKLS1 clusters with the non UV-inducible TaKSL4 rather than similarly regulated TaKSL2.

3. Discussion
The results reported here demonstrate that the KSL gene family in wheat has undergone expansion and functional diversification. This resembles the results described for the upstream CPS gene family in the accompanying report (Wu et al., submitted). Together, these results are consistent with physiological role(s) for the derived labdane-related diterpenoids. In rice, which has undergone similar CPS and KSL gene family expansion, the resulting natural products are thought to serve as inducible phytoalexins against fungal pathogens, constitutive phytoanticipans against bacterial infection, and/or as allelochemicals (Peters, 2006; Toyomasu, 2008). Given the similar transcriptional induction of TaKSL1, TaKSL2, and TaKSL5-1 and TaKSL5-2 as observed with the phytoalexin relevant OsKSL, it seems likely that the derived diterpenoids will serve as phytoalexins in wheat. On the other hand, the conservation of HvKS and TaKSL6 with OsKS1 strongly indicates that these are involved in gibberellin biosynthesis. Finally, it is possible that the noninducible TaKSL3 and TaKSL4 might serve in phytoanticipan biosynthesis.

The dual normal/syn-CPP reactivity observed here with TaKSL1 and TaKSL4 provides some insight into the underlying enzymatic catalysis, as the products resulting from reaction with these alternative substrates exhibit certain similarities (Fig. 5). Specifically, TaKSL1 produces a pimaradiene with β-methyl at C13 and C7,8-double bond resulting from deprotonation at C7 of the relevant isopimara-15-en-8-y1+ intermediate from both normal and syn-CPP. TaKSL4 produces a pimaradiene with α-methyl at C13 and double bond involving C9 from both substrates. However, when reacting with syn-CPP, this entails a hydride shift from C9 to C8, enabling deprotonation at C11 to form the observed C9,11-double bond, rather than the direct deprotonation at C9 of the initially formed pimara-15-en-8-yl+ intermediate observed upon reaction with normal CPP. The shared C13 configuration of the
pimaradienes resulting from reaction with these alternative substrates further implies that both TaKSL1 and TaKSL4 bind normal and syn-CPP in a similar conformation (Fig. 6). On the other hand, the production of abietadiene (10) from normal CPP by TaKSL2 prevents insight into the configuration of C13 in the relevant pimarenyl+ intermediates, which then cannot be compared to the ent-CPP conformation implied by the observed production of ent-pimara-8(14),15-diene (9).

The dual reactivity of TaKSL2 with the normal and ent-CPP found in wheat is similar to the ability of OsKSL10 to react with the syn- and ent-CPP found in rice. However, while the dual normal/syn-CPP reactivity of TaKSL1 and TaKSL4 found here is analogous to that reported for OsKSL4 and OsKSL11 (Morrone et al., in press), these do not seem to be physiologically relevant. Specifically, because rice does not produce normal CPP and, as described in the accompanying report (Wu et al., submitted), wheat does not seem to produce syn-CPP. Nevertheless, such latent plasticity presumably enables facile evolution of diterpenoid metabolism, as changes in CPP stereochemistry arising from functional diversification of the upstream CPS, would be readily accommodated by these promiscuous KSL. In addition, downstream enzymes such as cytochromes P450 exhibit similar promiscuity (Wang et al., resubmitted). Such broader metabolic plasticity enables the immediate appearance of multi-step biosynthetic pathways from changes in an upstream enzyme. Indeed, such changes in CPS stereochemistry appears to have been directly relevant in the cereal crop plant family given the differing production of normal or syn-CPP by CPS homologs from wheat and rice described in the accompanying report (Wu et al., submitted). On the other hand, it may be these evolutionary changes in CPS activity that led to the observed dual reactivity with some of the rice and wheat KSL, perhaps reflecting
promiscuous intermediate stages in the evolution of altered substrate specificity, a concept that has been more generally postulated (Tawfik, 2010).

4. Conclusion

The results presented here demonstrate that wheat contains an expanded and functionally diverse family of KSL. While the physiological roles of the ensuing labdane-related diterpenoids largely remain unclear, this gene family has undergone continued evolution in both wheat and rice since the separation of these related cereal crop plants, suggesting relevance for these natural products. Given the similar inducible transcriptional regulation of unrelated KSL from wheat and rice, it seems likely that wheat uses labdane-related diterpenoids in plant defense, analogous to their roles in rice. Regardless of physiological role, the biochemical activity exhibited by the KSL characterized here significantly expands the diterpenoid metabolic repertoire of wheat. In addition, the dual substrate reactivity demonstrated here for some of the wheat KSL echoes previous findings with rice KSL family members (Morrone et al., in press), and such enzymatic promiscuity presumably enabled the evolution of expanded labdane-related diterpenoid metabolism observed in the cereal crop plant family.

5. Experimental

5.1 General

Unless otherwise noted, all chemical reagents were purchased from Fisher Scientific (Loughborough, Leicestershire, UK), and molecular biology reagents from Invitrogen (Carlsbad, CA, USA). All recombinant expression was carried out with the OverExpress C41 strain of E. coli (Lucigen, Middleton, WI, USA). Gas chromatography with mass spectrometric detection (GC-MS) analyses were performed using a Varian (Palo Alto, CA,
USA) 3900 instrument with HP-5ms column and Saturn 2100 ion trap mass spectrometer in electron ionization (70 eV) mode. Samples (1 µL) were injected in splitless mode at 50 °C and, after a 3 min. hold, the temperature raised at 14 °C/min. to 300 °C, where it was held for 3 min. MS data was collected from m/z from 90 to 600, starting 12 min. after the injection until the end of the run. GC with flame ionization detection (FID) was carried out using an Agilent (Santa Clara, CA, USA) 6890 GC with HP-5ms column, and the same protocol utilized for GC-MS analyses. Bioinformatic sequence analyses were carried out either with the VectorNTI (Fig. 1) or CLC Sequence Viewer (Fig. 4) software packages.

5.2 Cloning

BLAST searches were carried out using the OsKSL as probes at the wheat genome database maintained by the J. Craig Venter Institute (http://www.jcvi.org/wheat/index.php). TaKSL1, 2, 3, 4, 5 and 6 correspond to the sequences of BE585476, CV775031, BE585481, CK205050/CA646242, DR740781/CJ593452, and BE419989. The corresponding full-length cDNA were then cloned from T. aestivum cv. Nourin-61-gou by rapid amplification of cDNA ends (RACE) and end-to-end RT-PCR using gene specific primers based on these EST sequences, much as described previously (Toyomasu et al., 1998).

Each KSL was transferred to the Gateway vector system via PCR amplification and directional topoisomerization insertion into pENTR/SD/D-TOPO, with the ensuing constructs verified by complete gene sequencing. These clones were subsequently transferred via directional recombination to the T7-based N-terminal GST fusion expression vector pDEST15.

5.3 Functional characterization via metabolic engineering
Functional characterization of HvKS and the TaKSL was accomplished by use of our previously described modular metabolic engineering system (Cyr et al., 2007). Briefly, class I labdane-related diterpene synthases such as TaKSL are co-expressed, from pDEST expression vectors, with pACYC-Duet (Novagen/EMD) derived plasmids that carry a GGPP synthase and CPS (pGGxC). These lead to production of the three most common stereoisomers of CPP, specifically pGGnC leads to production of normal CPP, pGGeC leads to production of ent-CPP, and pGGsC leads to production of syn-CPP. Accordingly, HvKS and the TaKSL were separately co-expressed with each of the three pGGxC vectors (i.e., in all possible pairings). The resulting recombinant bacteria were then analyzed as previously described (Morrone et al., 2009). Briefly, 50 mL NZY liquid media cultures were grown with shaking to $A_{600}$ ~ 0.6 at 37 °C, the temperature reduced to 16 °C for 1 hr prior to induction with IPTG (added to a final concentration of 0.5 mM), followed by continued fermentation at 16 °C for an additional ~72 hr. These cultures were then extracted with an equal volume of hexanes, dried under N$_2$, and resuspended in 100 μL fresh hexanes for GC-MS analysis, with product identification accomplished by comparison to authentic standards.

5.4 Analysis of inducible wheat labdane-related diterpene synthase gene

To measure TaKSL mRNA levels in response to UV irradiation, leaf sheaths were obtained from wheat plants that had been cultivated in a growth chamber for 2 weeks at 25°C and exposed to UV light for 15 min according to a previously described method (Otomo et al., 2004b), and then harvested 20 h or 40 h after irradiation. Total RNA was extracted from frozen samples using an RNAqueous column with Plant RNA Isolation Aid (Ambion), and cDNA was synthesized from 1-μg aliquots of total RNA by using a QuantiTect reverse transcription kit (Qiagen). Real-time QRT-PCR, using SYBR Green II, was carried out in a
TP800 thermal cycler (Takara). The mean values from two replicates were normalized using 18S rRNA as an internal control, as previously described (Sawada et al., 2008), with the primers listed in Table 1.

Acknowledgements

We thank Dr. Robert M. Coates (Univ. of Illinois) for graciously providing an authentic sample of synthetic beyer-15-ene. This study was generously supported by grants from the USDA-NIFA-AFRI (2008-35318-05027) and NIH (GM076324) to R.J.P.

References


Hillwig, M. L., Xu, M., Toyomasu, T., Tiernan, M. S., Gao, W., Cui, G., Huang, L., Peters, R. J., submitted. Lack of conservation in bi-domain diterpene synthases indicates complex evolutionary origins for plant terpene synthases. Plant J.


**Figures**

**Figure 1:** HvKS and TaKSL1-6 amino acid sequence alignment.
Figure 2: Transcriptional induction of wheat labdane-related diterpene synthase genes by UV-irradiation.
Figure 3: Identification of products formed by TaKS(L) via GC-MS based comparison to authentic standards.  (a-d) Chromatograph of the products formed by HvKS, TaKSL6, TaKSL5-1, and TaKSL5-2, respectively, from ent-CPP (2).  (e-i) Mass spectrum of HvKS and TaKSL6 products, as well as the major product of TaKSL5-1 and TaKSL5-2, from 2, with comparison to authentic ent-kaur-16-ene (RT = 16.81 min.).  (j-l) Mass spectrum of the minor product of TaKSL5-1 and TaKSL5-2 from 2, with comparison to authentic ent-beyer-15-ene (RT = 16.05 min.).  (m) Chromatograph of the product formed by TaKSL1 from normal CPP (3).  (n,o) Mass spectrum of TaKSL1 product from 3, with comparison to authentic isopimara-7,15-diene (RT = 16.49 min.).  (p) Chromatograph of the product formed by TaKSL4 from 3.  (q,r) Mass spectrum of TaKSL4 product from 3, with comparison to authentic pimara-8(9),15-diene (RT = 15.91 min.).  (s) Chromatograph of the product formed by TaKSL1 from syn-CPP (4).  (t,u) Mass spectrum of TaKSL1 product from 4, with comparison to authentic syn-isopimara-7,15-diene (RT = 15.82 min.).  (v) Chromatograph of the products formed by TaKSL4 from 4.  (w,x) Mass spectrum of TaKSL4 product from 4, with comparison to authentic syn-isopimara-9(11),15-diene (RT = 16.26 min.).
Figure 4: Molecular phylogenetic tree for characterized cereal KS(L).
Figure 5: Reactions catalyzed by HvKS and the various TaKSL. Also shown is generic initially catalyzed cyclization to a pimar-15-en-8-yl carbocationic intermediate, with carbon numbering indicated as mentioned in the text.
Figure 6: Comparison of the configuration of the alternative substrates in the active site of the dual normal/syn-CPP reactive TaKSL1 (A) and TaKSL4 (B).
### Table 1: Primers used in RT-PCR analysis of mRNA levels

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaKSL1</td>
<td>GCGGTTAACTCATTGCAGA</td>
<td>CCTCACTTTGACTCCCTCTTGA</td>
</tr>
<tr>
<td>TaKSL2</td>
<td>ATGTGGAGGAGGCATCTGC</td>
<td>GGCAACAACCTCAGCTCCAGG</td>
</tr>
<tr>
<td>TaKSL3</td>
<td>CTCTTGGCATCTGTGTGAATGG</td>
<td>GTTGAGGAGTCGGCAACAAG</td>
</tr>
<tr>
<td>TaKSL4</td>
<td>CGCTTACCTCATACGGGATG</td>
<td>CGTCTCTGATCCCTCTCA</td>
</tr>
<tr>
<td>TaKSL5</td>
<td>GATCAAGAGTGTCTGGACTTCA</td>
<td>CAGATGAACGGAGGGCTTCG</td>
</tr>
<tr>
<td>TaKSL6</td>
<td>ACTCACCCGATTTCGCT</td>
<td>GCGGTAGTTCAACCTTGCG</td>
</tr>
<tr>
<td>18S</td>
<td>GGAGCGATTTGTCTGGTTA</td>
<td>ATCTAAGGGCATACAGACC</td>
</tr>
</tbody>
</table>
Chapter V: Direct action of a single residue switch for diterpene synthase product outcome

Manuscript in preparation and to be submitted to Organic & Biomolecular Chemistry

Ke Zhou and Reuben J. Peters*

Abstract

Exchange of aliphatic and aliphatic hydroxyl side chains at a single amino acid position can have dramatic effect on product outcome in diterpene synthases. The location of the relevant residue differs slightly between various enzymes, and the effect of side chain chemistry has not been further explored. Here we show that the ability of these single residue changes to affect product outcome is specific for both active site location and side chain chemical composition, as well as further demonstrate a direct interaction between the relevant residue and carbocation intermediate.

Introduction

Terpene synthases create diverse hydrocarbon backbone structures from isoprenyl diphosphate precursors via carbocationic cyclization and rearrangement reactions that often proceed with remarkable regio- and stereochemical specificity. These enzymes typically carry out the committed step in their particular biosynthetic pathway. Hence, there has been significant interest in the mechanisms by which terpene synthases exert control over the catalyzed carbocationic reaction.

Recent studies have demonstrated that terpene synthases exhibit extreme plasticity, with small numbers of amino acid substitutions being sufficient to substantially alter product outcome. Among the most striking of these are reports on the ability of single residue
changes to dramatically shift the product outcome of terpene synthases involved in labdane-related diterpene biosynthesis (i.e., those catalyzing the production of further cyclized and/or rearranged diterpene olefins from labdadienyl/copalydiphosphate – CPP).\textsuperscript{2-5}

Spurred by identification of highly homologous functionally divergent paralogs from rice (\textit{Oryza sativa}),\textsuperscript{6,7} Thr substitution for a conserved Ile was found to be sufficient to “short circuit” the complex cyclization and rearrangement reaction catalyzed by \textit{ent}-kaurene synthases, leading to the production of \textit{ent}-pimaradiene instead (presumably via deprotonation of a mechanistically relevant pimarenyl\textsuperscript{+} intermediate).\textsuperscript{2} Conversely, Ile substitution for the corresponding Thr in a \textit{syn}-pimaradienesynthase was found to dramatically increase the complexity of the catalyzed reaction, resulting in predominant production of a rearranged tetracycle rather than tricyclic pimaradiene.\textsuperscript{5}

The corresponding position did not exhibit the same Ile/Thr variation in side chain composition in comparisons of labdane-related diterpene synthases acting on CPP of normal (1), rather than \textit{ent}- or \textit{syn}-, stereochemistry (Figure1). However, analogous functionally coupled variation was found in such TPS producing either the rearranged tricyclic abietadienes (2a-d) or isopimara-7, 15-diene (3a) at a nearby position four residues away. Specifically, Ala in abietadiene synthases and Ser in the only known isopimaradiene synthase.\textsuperscript{8} Exchange of residues at this position gave a dramatic shift in product output, much as observed with the \textit{ent}-kaurene synthases, although an additional change was required to induce production of the rearranged abietadienes by the native isopimaradiene synthase to.\textsuperscript{3,4}

Regardless of exact location, it has been hypothesized in all of these reports that the relevant residue directly interacts with the initial carbocation in the pimarenyl\textsuperscript{+} intermediate
formed by (tri)cyclization of the relevant CPP (Figure 2). Given the expected helical nature of the corresponding enzymatic sequence, the difference in position between the two previously identified single residue switches for diterpene synthase product outcome corresponds to one turn, or a change in active site location of ~4 Å. However, the ability of changes at the alternate position to modify product output has not been previously investigated, nor has the hypothesized direct interaction between the relevant residue and pimarenyl Intermediate been probed in any other way.

To investigate the specificity of the previously identified single residue product output switch, we returned to the model diterpene cyclase abietadiene synthase from *Abies grandis* (AgAS). In previous work we demonstrated that Ser substitution for Ala 723 in AgAS (A723S) led to a dramatic alteration in product outcome from >95% abietadienes to >95% (iso)pimaradienes (3a-b). Here we substituted Thr for Val 727 (AgAS:V727T), the position corresponding to the single residue switch operating in the ent-or syn-, rather than normal, CPP specific labdane-related diterpene synthases, and Cys for Ala 723 (AgAS:A723C), to more directly probe the hypothesized interaction.

**Results**

The product outcome mediated by both mutant AgAS was analyzed by incorporation into our previously reported modular metabolic engineering system. This characterization demonstrated that the V727T mutation did not alter product outcome relative to wild type AgAS, while the A723C mutation did have a relatively subtle, compared to the A723S mutation, but appreciable effect (Table 1). The strains engineered with the mutants produced similar amounts of diterpenes as found with wild type AgAS (within 3-fold).
However, more detailed steady-state kinetic analysis demonstrated that the mutations had more significant deleterious effects (>5-fold) on catalytic activity in vitro (Table 2).\textsuperscript{12,13}

Of particular interest is the A723C mutant, which we had hypothesized, might be alkylated by formation of a covalent bond between the introduced thiol and any neighboring carbocation reaction intermediate. This mutant exhibited a 300-fold reduction in specific activity (kcat/K_M) when directly fed labeled CPP (1). However, assays fed unlabeled 1 prior to kinetic assessment using labeled 1 demonstrated a rapid loss of activity with the A723C mutant, but not the other AgAS variants, consistent with alkylation (Figure 3). This was directly demonstrated by the molecular weight comparison via QTOF for the A723C incubated with and without substrate GGPP. A mass shift difference of 272 Da between two samples indicates one geranylgeranylation (Figure 4). Further LC-MS assay for digested samples confirmed it by detecting unmodified Y659-K698 in the non-incubated sample, and +272 on the Y659-K698 peptide in the incubated sample (Figure 5). Cys 675 inside this peptide is actually the one introduced into 723 position because the AgAS studied here is a truncated version (d83) in which the transit peptide was trimmed, and pDEST17 vector results in a 25 amino acid residue linker between the 6 x His tag and the cloned protein, as well as 4 more in the front of 6 x His tag.

**Discussion**

The ability of the A723S mutation to alter product outcome is remarkable. Wild type AgAS is quite specific for production of abietane type relative to pimarane type hydrocarbon backbone structures, exhibiting a specificity factor of 50 for abietane type diterpene production (Σ2a-d/Σ3a-b). By contrast, AgAS:A723S specifically produces pimarane type diterpenes, exhibiting a corresponding specificity factor of ~24 (Σ3a-b/Σ2a-d). This
represents a >1,000-fold increase in relative product type ratio (Table 1). By this measure, the change in product outcome mediated by Cys substitution at the same position is much less dramatic. AgAS:A723C exhibits only a slight preference for production of abietadienes relative to pimaradienes (specificity factor ~3), representing a relatively small (<20-fold) change in relative product ratio.

Examination of the cyclization and rearrangement of 1 catalyzed by AgAS offers a slightly different, and perhaps mechanistically more relevant view (Scheme 1). Production of abietadiene is known to proceed through an isopimar-15-en-8-yl\(^{+}\) intermediate (4a\(^{+}\)).\(^{14}\) This can either undergo deprotonation to a pimaradiene (3a-b) or intramolecular proton transfer to create the isopimar-8(14)en-15-yl\(^{+}\) intermediate (4b\(^{+}\)) from which methyl migration and subsequent deprotonation produces the abietadienes(2a-d). Thus, the competing deprotonation (4a\(^{+}\)→3a-b) and intramolecular proton transfer (4a\(^{+}\)→4b\(^{+}\)) represents the bifurcation point between production of pimaradienes versus abietadienes. In wild type AgAS, >98% of the initially formed 4a\(^{+}\) undergoes proton transfer to form 4b\(^{+}\) and, subsequently, 2a-d. In AgAS:A723C this drops to ~73%, and in AgAS:A723S to ≤4%. While somewhat less dramatic, even this arguably more realistic view of the relative effects of Ser versus Cys substitution for the Ala normally found at this position in AgAS demonstrates that introduction of a hydroxyl group has a significantly greater effect than introduction of a thiol.

We have previously hypothesized that the amino acid side chain at position 723 of AgAS interacts with the 8-ylcarbocation. In particular, that this residue directly affects the kinetic competition between the competing reaction channels 4a\(^{+}\)→3a-b and 4a\(^{+}\)→4b\(^{+}\), with transient stabilization of the 8-yl of 4a\(^{+}\) by an aliphatic hydroxyl group enabling
deprotonation to form pimaradienes. The results we report here bolster this hypothesis. Introduction of a nearby aliphatic hydroxyl (V727T), despite little change in steric volume, does not affect product outcome. Even substitution of a thiol for the critical hydroxyl (A723C versus A723S) is sufficient to significantly reduce pimaradiene formation. The alkylation indicates that the thiol or thiolate on introduced Cys side chain is close enough to $4a^+$-yl, supporting a possible interaction between the side chain functional group at this specific position and this specific carbocation intermediate. Thus, for those escaped from alkylation, the lower dipole moment of thiols relative to hydroxyls presumably underlies the reduced ability of Cys substitution to stabilize $4a^+$ sufficiently for deprotonation to occur. This specifically implies a direct polar interaction between the $4a^+$-yl and relevant side chain functional group, with the higher electronegativity of a hydroxyl oxygen providing a stronger interaction than a thiol sulfur, leading to the observed significant increase in deprotonation ($4a^+ \rightarrow 3a-b$) relative to competing intramolecular proton transfer ($4a^+ \rightarrow 4b^+$).

**Conclusion**

In conclusion, we have demonstrated here that the single residue switch for product outcome in labdane-related diterpene synthases is specific with respect to both active site location (AgAS:A723S versus AgAS:V727T) and side chain chemical composition (AgAS:A723S versus AgAS:A723C). Furthermore, our results support the previously suggested mechanism wherein the side chain of the residue at the switch position directly interacts with the $8$-yl carbocation of the initially formed pimaren vinyl intermediate, with the transient stabilization provided by a polar group, particular more electronegative hydroxyl oxygen, providing sufficient time for deprotonation to occur.
Experimental

General procedures

Unlabeled [E,E,E]-geranylgeranyl pyrophosphate (GGPP) was purchased from Isoprenoids, LC (Tampa, FL). $[^{3}\text{H}]$GGPP (1) was purchased from American Radiolabeled Chemicals (St. Louis, MO), Ni-NTA His-bind matrix was purchased from Novagen (Madison, WI), and molecular biology reagents were purchased from Invitrogen (Carlsbad, CA). Unless otherwise noted, all other chemicals were purchased from Fisher Scientific (Loughborough, Leicestershire, UK).

Recombinant constructs

The recombinant pseudo-mature AgAS gene has been previously described, and transferred into the Gateway (Invitrogen) vector system via directional topoisomerization into pENTR/SD/D-TOPO. Site-directed mutagenesis was carried out via PCR amplification of the pENTR/AgAS construct with overlapping mutagenic primers, and the mutant genes verified by complete sequencing. The resulting wild type and mutant genes were then transferred via directional recombination to the T7-promoter and N-terminal 6 x His fusion expression vector pDEST17.

Enzyme expression and purification

Transformed bacterial cells (C41) including pDEST17/AgAS (wild type or one of the various mutants) were grown and harvested like previous description. AgAS was purified using Ni-NTA His-bind resin. The procedure of purification has been previously described.

Enzymatic analysis

Kinetic analysis was carried out using the D621A mutant of AgAS, which abolishes class I
activity, to convert $[^1\text{H}]$GGPP to $[^1\text{H}]$copalyl pyrophosphate (CPP). In brief, 1 ml reactions in assay buffer (50 mM Hepes, pH 7.2, 0.1 mM MgCl2, 5% glycerol, 5 mM DTT, 100 mM KCl and 0.1 mg/ml a-casein) containing 50 uM $[^1\text{H}]$GGPP and 1 uM AgAS:D621A were run for 2 h at 30 °C, which is sufficient to completely convert this amount of unlabeled GGPP to CPP. Kinetic assays were then performed similar to the method described previously (Peters et al., 2000). In brief, duplicate 1 ml reactions in assay buffer containing 3 nM wild type AgAS were initiated by the addition of $[^1\text{H}]$CPP, allowed to react for 1 min at 30 °C, then stopped by the addition of KOH to 0.2 M and EDTA to 150 mM. For analysis of the mutants, the enzymatic concentration was increased to 6, 5, or 30 nM, and incubation times increased to 3, 2, or 30 min for V727T, A723S, or A723C, respectively. The produced diterpenes were then extracted by hexane, the pooled extract passed over a short silica gel column, and product formation assessed by scintillation counting. The resulting data was analyzed using Kaleidagraph (Synergy, Reading, PA).

Diterpene products were analyzed by co-expressing AgAS (wild type or one of the various mutants) with GGPP synthase in Escherichia coli, as previously described. Briefly, AgAS in pDEST17 expression vectors were co-expressed with pACYC-Duet (Novagen/EMD) derived plasmids (pGG), where GGPP synthases was carried on. In the following, C41 cells were co-transformed with pDEST17 expression vectors carrying AgAS genes (WT or mutants), along with pGG vectors. Growing of the recombinant C41 cells followed the procedure like previously described. The resulting organic solvent extract of these recombinant cells were analyzed by gas chromatography with mass spectral detection (GC–MS), which was performed with the same method as described previously.
**Incubation assay with non-radioactive-labeled CPP**

0.1 uM purified AgAS:A723C was incubated with non-radioactive-labeled normal CPP in the assay buffer for the time course with 0, 10, 20, 30 sec., 1, 2, 5, 10, 30, 60 min. Followed by the radioactive assay in the assay buffer containing 5 uM $^3$H-CPP for one hour. The procedure of enzymatic assay is same as description above.

**HPLCchip-QTOF and LC-MS measurement for substrate-incubated AgAS:A723C**

1 uM AgAS: A723C was incubated in the buffer (50 mM Hepes, pH 7.2, 7.5 mM MgCl2, 1% glycerol, 5 mM DTT, 100 mM KCl) with or without 100 uM GGPP at 30 °C for 2 hours. The samples were shipped to Proteomics and Mass Spectrometry Facility at Donald Danforth Plant Science Center. The samples for QTOF were firstly zip-tipped. The samples for LC-MS were firstly digested in-solution.

**Acknowledgment.** This work was supported by NIH grant GM076324 to R.J.P.

**References**

9. Mutants were created by standard PCR based protocols, and verified by complete sequencing, much as previously described (refs. 2, 3, & 5).
11. Products were identified by comparison of retention time and mass spectra to
that of authentic standards by GC-MS. Relative product amounts were determined by comparing peak areas from total ion chromatograms.

(12) Kinetic measurements were carried out as previously described (ref. 13).


Table 1: Normalized (%) Product Amounts of AgAS wild type and mutants.

<table>
<thead>
<tr>
<th>AgAS:</th>
<th>2a</th>
<th>2b</th>
<th>2c</th>
<th>2d</th>
<th>3a</th>
<th>3b</th>
<th>( \Sigma 2a-d/\Sigma 3a-b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>48</td>
<td>25</td>
<td>21</td>
<td>4</td>
<td>-</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>A723S</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>76</td>
<td>20</td>
<td>0.04</td>
</tr>
<tr>
<td>A723C</td>
<td>36</td>
<td>18</td>
<td>15</td>
<td>4</td>
<td>11</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>A727T</td>
<td>48</td>
<td>24</td>
<td>21</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 2: Steady-state kinetic constants of AgAS wild type and mutants

<table>
<thead>
<tr>
<th>AgAS:</th>
<th>( k_{cat} ) (s(^{-1}))</th>
<th>( KM ) (( \mu )M)</th>
<th>( k_{cat}/KM ) ( (\times\ 105\ M^{-1}\ s^{-1}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.4 ± 0.1</td>
<td>1.2 ± 0.3</td>
<td>3</td>
</tr>
<tr>
<td>A723S</td>
<td>0.03 ± 0.01</td>
<td>0.5 ± 0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>A723C</td>
<td>0.004</td>
<td>0.3 ± 0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>A727T</td>
<td>0.07 ± 0.01</td>
<td>3.7 ± 0.4</td>
<td>0.2</td>
</tr>
</tbody>
</table>
Figure 1: Alignment of crucial residue switch for catalytic specificity for selected diterpene synthases #: the position of key single residue in olefin synthases; *: the position in OsKSL synthases.

\[
\begin{array}{ll}
Ent- & OsKSL5i (657) \text{DSFAVGP} \text{IITS} \\
Ent- & OsKSL5j (657) \text{DSFALGPT} \text{TITS} \\
Syn- & OsKSL4 (689) \text{VTFALGPT} \text{ILI} \\
\# & * \\
Normal- & AgAS (720) \text{VSIALGT} \text{VVLI} \\
Normal- & PaAS (710) \text{VSIALGTVVLI} \\
Normal- & PaPS (718) \text{VSIISLGT} \text{LVLI} \\
\end{array}
\]
Figure 2: Single residue switch is proposed close to the pimarenyl\textsuperscript{+} intermediate.
Figure 3: Rapid loss of activity with the A723C mutant during the incubation with CPP substrate.

V(uM/S)

Incubated time with unlabeled CPP (Sec)
Figure 4: Intact mass using HPLCchip-QTOF. S1: AgAS:A723C did not incubated with GGPP. S2: AgAS:A723C incubated with GGPP.

<table>
<thead>
<tr>
<th>Compound Label</th>
<th>Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>94976.0693</td>
</tr>
<tr>
<td>S2</td>
<td>95248.1743</td>
</tr>
</tbody>
</table>

Mass Shift: 272 Da
Figure 5: Digested AgAS:A723C followed by LC-MS. Sample 1-1: AgAS:A723C did not incubated with GGPP. Sample 1-2: AgAS:A723C incubated with GGPP.
Sample 1-1 - unmodified: Peptide Y659-K698 (containing Cys 675): MW theoretical: 4321.9 (4+ charge state = 1081.48 m/z) MW experimental: 4321.2 (4+ charge state = 1081.3 m/z)

Sample 2-1 – modified: Peptide Y659-K698 (containing Cys 675) + 272 Da: MW theoretical: 4593.9 (4+ charge state = 1149.48 m/z) MW experimental: 4593.8 (4+ charge state = 1149.46 m/z)
Scheme 1: AgAS catalyzed reaction
Chapter VI: Electrostatic effects on (di)terpene synthase product outcome

This review was on-line published in and reprinted from Chemical Communications. 2011 Feb 8. [Epub ahead of print]

Ke Zhou and Reuben J. Peters

Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

Abstract

Terpene synthases catalyze complex reactions, often forming multiple chiral centers in cyclized olefin products from acyclic allylic diphosphate precursors, yet have been suggested to largely control their reactions via steric effects, serving as templates. However, recent results highlight electrostatic effects also exerted by these enzymes. Perhaps not surprisingly, the pyrophosphate co-product released in the initiating and rate-limiting chemical step provides an obvious counter-ion that may steer carbocation migration towards itself. This is emphasized by the striking effects of a recently uncovered single residue switch for diterpene synthase product outcome, whereby substitution of hydroxyl residues for particular aliphatic residues has been shown to be sufficient to “short-circuit” complex cyclization and/or rearrangement reactions, with the converse change further found to be sufficient to increase reaction complexity. The mechanistic hypothesis for the observed effects is hydroxyl dipole stabilization of the specific carbocation formed by initial cyclization, enabling deprotonation of this early intermediate, whereas the lack of such stabilization (i.e. in the presence of an aliphatic side chain) leads to carbocation migration towards the pyrophosphate co-product, resulting in a more complex reaction. This is further consistent with the greater synergy exhibited between pyrophosphate and aza-analogs of late, relative to early, stage carbocation.
intermediates, and crystallographic analysis of the monoterpane cyclase bornyl diphosphate synthase wherein mechanistically non-relevant counter-ion pairing between aza-analogs of early stage carbocation intermediates and pyrophosphate is observed. Thus, (di)terpene synthases seem to mediate specific reaction outcomes, at least in part, by providing electrostatic effects to counteract those exerted by the pyrophosphate co-product.

Introduction

Terpenoids are the most structurally diverse class of natural products, with over 50,000 already known. This chemical diversity is underlaid by manyfold hydrocarbon backbone structures formed by the cyclization and/or rearrangement of acyclic precursors catalyzed by terpene synthases. These precursors arise from coupling of the universal 5-carbon isoprenoid precursors isopentenyl diphosphate and dimethylallyl diphosphate into longer chain polyisoprenoid diphosphates. Of particular interest here are the 10-carbon monoterpane precursor geranyl diphosphate (GPP), 15-carbon sesquiterpane precursor farnesyl diphosphate (FPP), and 20-carbon diterpene precursor geranylgeranyl diphosphate (GGPP). Typically, the allylic diphosphate ester bond of these acyclic precursors is then ionized by a terpene synthase to initiate electrophilic reactions that transform these into the observed manyfold cyclized and/or rearranged hydrocarbon backone structures. Although GGPP may first undergo a separate protonation-initiated bicyclization reaction catalyzed by mechanistically distinct (class II) diterpene cyclases, this leaves intact the allylic diphosphate ester bond in the resulting labdane-related diterpenoid biosynthetic intermediate, enabling subsequent ionization and further transformation by members of the more typical (class I) terpene synthase
enzyme family that are the focus of this review.

**Terpene synthases**

As suggested by their enzymatic classification (EC 4.2.3.x), terpene synthases use lysis/ionization of the allylic diphosphate ester bond to drive carbon bond formation. Despite exhibiting very little to no overall sequence similarity, crystallographic investigations have revealed that terpene synthases from plants, fungi, and bacteria share protein structure homology.² Specifically, catalysis occurs in an analogous α-helical bundle tertiary assembly that has been termed the class I terpene cyclase fold.⁴ Catalysis further relies on a trinuclear cluster of divalent metal ion co-factors (generally magnesium), which are coordinated by two binding motifs that provide the only broadly conserved sequence features between microbial and plant terpene synthases,⁵,⁶ and which are further coordinated to the substrate diphosphate moiety to enable initiating ionization (Figure 1). Nevertheless, the observed enzymatic structural and mechanistic similarities indicates common origins for all terpene synthases, making the observations noted here broadly applicable.

**Basic catalytic mechanism**

Terpene synthases typically catalyze cyclization reactions, which are mediated by intramolecular carbon-carbon double bond addition to carbocation intermediates. Ionization of the diphosphate ester generates an allylic carbocation that undergoes concerted addition to another π bond within the same substrate molecule, forming a carbocation at a different carbon center, in an $S_N'$ reaction. Such initial cyclization can be followed by further cyclization and/or rearrangements mediated by proton, hydride and/or
methyl shifts with correlated carbocation migration. The resulting series of carbocations is then terminated by quenching of the final such intermediate, most typically by deprotonation to yield an olefin, although this also can occur after addition of water to yield a hydroxylated or cyclic ether containing product, or via recapture of the ionized pyrophosphate to generate a diphosphate.\footnote{8}

**Substrate folding – the template enzyme model**

Implicit in this $S_N'$ reaction mechanism is the necessary folding of the substrate by the enzyme to bring together the incipient allylic carbocation and carbon-carbon double bond to be joined (Figure 2).\footnote{9} In addition, the proximity of this $\pi$ bond presumably provides anchimeric assistance for the initiating ionization of the allylic diphosphate ester bond, which is viewed as the rate limiting chemical step, although enzymatic turnover seems to be limited by product release.\footnote{10, 11} This has led to the view that terpene synthases may dictate product outcome in large part by simply providing a product-like template into which their substrate is folded prior to triggering the relevant carbocation cascade by ionization.\footnote{12} Indeed, it has been observed that the active site of terpene synthases in the “closed” conformation, but in the absence of any substrate or reaction intermediate analog, are nevertheless distinctly “product-like” in at least some cases.\footnote{13} Thus, the template model implies that terpene synthases dictate reaction outcome by restricting their substrate and discrete reaction intermediates to a subset of the possible conformations.

**Kinetic rather than thermodynamic control**

Intriguingly, multiple examples exist of terpene synthase crystal structures wherein substrate or reaction intermediate analogs are bound in non-catalytically relevant
conformations in the active site,\textsuperscript{7, 14-19} although examples also exist wherein relevant conformations are observed, even of the same enzymes with different analogs.\textsuperscript{7, 17-19} Notably, the relevance of the observed substrate/intermediate analog configuration to catalytic mechanisms seems to depend in large part on how closely the analog mimics the reaction product.\textsuperscript{2} For example, monoterpene cyclases are considered to proceed via initial rearrangement of GPP to the tertiary linalyl diphosphate (LPP), and co-crystal structures of limonene synthase with the inert 2-fluoro analogs of each reveal that exogenous 2F-LPP, but not 2F-GPP (even though it is enzymatically converted to 2F-LPP), is bound in a catalytically relevant conformation.\textsuperscript{18} Particularly informative is the series of tertiary crystal structures of bornyl diphosphate synthase (BPS) complexed with inorganic pyrophosphate and/or aza-analogs, mimicking various carbocations of the BPS catalyzed reaction (Figure 3).\textsuperscript{7} In both cases where there is a separate pyrophosphate group, the aza moiety was ion-paired with the pyrophosphate, whose position was essentially invariant (as fixed by its interactions with the enzyme and bound divalent metal ion co-factors). Only in the case of the aza-bornane analog of the final carbocation intermediate does the co-crystal conformation appear to be catalytically relevant. Indeed, the 7-aza-7,8-dihydrolimonene analog is clearly bound “backwards” to enable aza-pyrophosphate ion-pairing, indicating that this is the thermodynamically favored binding mode for such a carbocation (Figure 3). Thus, it seems clear that terpene synthases exert kinetic, rather than thermodynamic, control over the catalyzed reactions — i.e., intermediates in the chemical reaction do not necessarily reorient to the most energetically favorable conformation, as they are only transiently present and their movement is restricted by the enzyme active site.\textsuperscript{15}
A role for the pyrophosphate anion co-product

The role of the substrate diphosphate moiety extends beyond that of a simple leaving group. As noted above, it has been demonstrated that the first step in mono-, as well as many sesqui-, terpene cyclization reactions is diphosphate migration from C-1 to the tertiary C-3 position, with concurrent double bond shift to enable C-1 to C-6 bond formation, which is otherwise prevented by the trans configuration of the original C-2,3 double bond (Figure 3).

Furthermore, inorganic pyrophosphate, corresponding to the anionic co-product released by initiating ionization, is able to drive terpene synthase active site closure.\(^7\), \(^13\), \(^15\), \(^20\) Given that terpene synthase enzymatic turnover is limited by product release,\(^10\), \(^11\) this indicates that the pyrophosphate co-product is retained in the active site during the electrophilic reaction cycle. Indeed, it has been suggested that the pyrophosphate co-product is tightly bound and may serve as a general acid/base during terpene synthase reactions without becoming reattached itself.\(^13\), \(^21\) Further consistent with such tight binding is the stereospecificity of bornyl diphosphate formation by BPS, which reattaches the bornyl cation to the same oxygen of the diphosphate involved in the original diphosphate ester bond of the GPP substrate, indicating that the pyrophosphate anion remains in a fixed orientation during the catalyzed reaction.\(^22\) Nevertheless, BPS also produces a significant amount of various olefinic products as well,\(^23\) with the production of bornyl diphosphate seeming to depend on subtle changes in active site conformation rather than electrostatic effects,\(^8\) consistent with a role for the pyrophosphate anion co-product in acting as a general base to terminate cyclization.

Obviously, the ionized pyrophosphate and carbocation intermediates of the released
olefin must be separated to prevent recapture (except for the bornyl cation in the case of the BPS catalyzed reaction). Nevertheless, we have suggested that the pyrophosphate co-product also may affect the reaction outcome by providing a counter-ion that influences carbocation migration. This was first based on our extensive studies of the model diterpene cyclase, abietadiene synthase (AS).\textsuperscript{24-30} Specifically, despite the inability of inorganic pyrophosphate to inhibit AS alone, which is unlike mono- and sesequi- terpene synthases,\textsuperscript{25} it strongly potentiates inhibition of AS by an aza-analog of a late stage intermediate in the catalyzed reaction, increasing affinity by over three orders of magnitude.\textsuperscript{27} This is indicative of strong ion-pairing, and is a much stronger effect than observed with other such carbocation mimic analogs,\textsuperscript{31-33} including earlier stage reaction intermediate analogs with AS (R.J.P., H.-J. Lee, R.B. Croteau, and R.M. Coates; unpublished results). Further evidence for the ability of the pyrophosphate anion co-product to influence diterpene product outcome is presented here below.

**A single residue switch for diterpene synthase product outcome**

It has long been supposed that terpene synthases would exhibit some ability to stabilize at least certain carbocation reaction intermediates. However, while it has been suggested that terpene synthase exert such electrostatic effects, via both carbocation-quadrupole interactions with the ring $\pi$ electrons of aromatic side chains\textsuperscript{34} and/or carbocation interactions with fixed and protected dipoles in the enzyme active site,\textsuperscript{2} until recently there was no direct evidence for such interactions.

Below we present recent studies from our group that provide strong evidence for a electrostatic role of a hydroxyl dipole in a wide range of diterpene synthase catalyzed
reactions, and further support a role for the pyrophosphate anion co-product in steering carbocation migration during the olefin cyclization-rearrangement reaction.

**Labdane-related diterpene synthases**

As noted in the Introduction, in diterpene biosynthesis GGPP is often first cyclized in a separate protonation-initiated reaction to a bicyclic intermediate such as labdadienyl/copalyl diphosphate (CPP, Figure 4). This biosynthetic intermediate is then further cyclized and/or rearranged by a more typical (class I) terpene synthase. The derived large group of natural products (~7,000 known) has been termed the labdane-related diterpenoids. Notably, the relevant (labdane-related class I) diterpene synthases are generally specific for such bicyclic diphosphates, exhibiting much less reactivity with more typical acyclic substrates such as GGPP. The presence of the bicycle in these imparts rigidity relative to acyclic substrates such as GGPP, so that accommodating changes in substate conformation necessary for altered product outcome requires significant remodelling of the active site. Hence, we have hypothesized that these labdane-related diterpene synthases can then serve as a model system for analysis of terpene synthase substrate and product specificity.

**Discovery**

Because terpene synthases are generally conserved by taxonomic origin rather than biochemical function, we undertook a functional genomics investigation of rice (*Oryza sativa*). In particular, because rice was known to make a number of labdane-related diterpenoid natural products, the corresponding multiple diterpene synthase activities had been demonstrated, and the agronomic importance of this staple cereal crop plant had led to early sequencing of its genome. This led to parallel investigations by our own
group and a consortium in Japan. The Japanese studies were focused on the rice subspecies *japonica* (cv. Nipponbare),\(^43-46\) while we largely worked with the *indica* subspecies (cv. IR24).\(^35-37\) Intriguingly, one functional difference between the labdane-related diterpene synthase arsenal of these subspecies became apparent,\(^37\) which proved to be extremely informative.

In particular, depending on subspecies origin (*indica* or *japonica*), orthologs of one of the rice diterpene synthases produced either tetracyclic *ent*-isokaur-15-ene or tricyclic *ent*-pimara-8(14),15-diene, respectively. Production of *ent*-pimaradiene represents deprotonation of the presumed *ent*-pimar-15-en-8-yl\(^+\) intermediate in the cyclization of *ent*-CPP to kaurane type diterpenes (Figure 5). This pair of functionally distinct subspecies orthologs share 98% identity at the amino acid (aa) sequence level, and there are only three differences in the active site.\(^37\) It was then possible to map their functional difference to a single residue, alteration of which was sufficient to essentially dictate product outcome, with the presence of a Thr at this position leading to production of *ent*-pimaradiene while an Ile leads to *ent*-isokaurene production,\(^47\) with similar *ent*-isokaurene production resulting from the presence of Val, which is more closely isosteric to Thr.\(^48\)

Furthermore, rice contains another, closely related *ent*-isokaurene synthase that shares 89% aa sequence identity with the functionally divergent orthologs, and also contains an Ile at this position. Similarly, Thr substitution then converts this enzyme to the production of *ent*-pimaradiene as well. Strikingly, this position is conserved as Ile in the *ent*-kaur-16-ene synthases found in all higher plants for gibberellin phytohormone biosynthesis and, despite sharing only 41-52% aa sequence identity with the rice *ent*
isokaurene synthases, Thr substitution had a very similar effect in two disparate ent-kaurene synthases, “short-circuiting” the more complex cyclization of ent-CPP to ent-kaurene to production of the “simpler” ent-pimaradiene. Given the highly variable context (i.e., divergent enzymatic sequences) in which this simple Ile to Thr change was found to “switch” product outcome, it seems unlikely that the introduced hydroxyl group is sufficiently activated to act as a general base to directly deprotonate the pimarenyl\(^+\) intermediate.

Accordingly, the ability of the inert aliphatic Ile to enable a more “complex” reaction mechanism (i.e. further cyclization and rearrangement), while introduction of a hydroxyl dipole short-circuits this, seems counter-intuitive. Specifically, this implies that there must be some other effect promoting further cyclization of the ent-pimarenyl\(^+\) intermediate, despite the accompanying, energetically unfavorable tertiary to secondary carbocation transition. Notably, studies with an aza-analog of the beyeran-16-yl secondary carbocation intermediate initially formed upon tetracyclization demonstrated that inhibition of ent-kaurane synthase by this 16-aza-ent-beyerane was strongly potentiated by inorganic pyrophosphate, suggesting strong ent-beyeranyl\(^+\)/pyrophosphate ion pairing in this cyclization-rearrangement reaction. However, this does not appear to lead to covalent bond formation, as the corresponding ent-beyeranyl diphosphate is not a substrate for ent-kaurene synthase, although it can be shown to bind in the active site as a competitive inhibitor. From these results, we hypothesized that the side chain of this product switch residue is adjacent to the ent-pimar-15-en-8-yl carbocation, such that the presence of a hydroxyl dipole stabilizes this intermediate long enough for deprotonation to occur, while the lack of such stabilization (i.e. in the presence of an aliphatic residue)
enables the tightly bound pyrophosphate anion co-product to steer carbocation migration towards itself. In particular, this leads to transient formation of an ent-beyeranyl\(^+\)/pyrophosphate ion pair, which may represent the transition state rather than a discrete intermediate\(^{50}\) with subsequent rearrangement to the more stable tertiary ent-kauran-16-yl\(^+\), which is still adjacent to the pyrophosphate counter-ion, prior to concluding deprotonation.

**Extension**

The hypothesized role for the pyrophosphate anion co-product in driving ent-kaurene synthase product formation is reminiscent of that we had previously advanced for abietadiene synthase (AS)\(^{27}\). In the AS catalyzed class I reaction, cyclization of CPP to isopimar-15-en-8-yl\(^+\) is followed by an intramolecular 1,4-proton transfer\(^{24,27,51}\), which forms a secondary isopimar-8(14)-en-15-yl\(^+\) that then undergoes a 1,2-methyl shift to the tertiary abieta-8(14)-en-13-yl\(^+\) prior to deprotonation (Figure 6), with the carbocation in the later two intermediates closer to the pyrophosphate anion co-product than that in the initial isopimar-15-en-8-yl\(^+\) (Figure 7). As noted above, our initial suggestion that the pyrophosphate anion co-product has a role in driving production of the rearranged abietadiene tricycles arose from the strong potentiating effect of inorganic pyrophosphate on binding of the 15-aza-isopimarene analog of the secondary isopimar-8-en-15-yl\(^+\) intermediate to AS\(^{27}\), while 14-aza-isopimarene, mimicking the initial tertiary isopimar-15-en-8-yl\(^+\), exhibits much less synergy with pyrophosphate (R.J.P., H.-J. Lee, R.B. Croteau, and R.M. Coates; unpublished results). While the residues at the position corresponding to the previously discovered switch for ent-kaurene synthase product outcome were conserved as aliphatic side chains in closely related abietadiene and
isopimara-7,15-diene synthases, we noted a similar hydroxyl/aliphatic conservation pattern four residues prior. This residue also is located in the active site (Figure 7), one turn of the helix away. Gratifyingly, substitution of the corresponding Ala in AS with Ser led to “short-circuiting” of the usual AS catalyzed cyclization-rearrangement reaction, with the resulting mutant enzyme producing essentially only isopimara-7,15-diene. This is analogous to the results with ent-(iso)kaurene synthases and resulting hypothetical mechanism described above, as well as our previously advanced hypothesis for the role of the pyrophosphate anion co-product in steering carbocation migration in the AS catalyzed reaction. Specifically, in the absence of any counteracting stabilizing influence (i.e. interaction with a hydroxyl dipole), the tightly bound pyrophosphate anion co-product steers carbocation migration towards itself. In this case, via intramolecular 1,4-proton shift to create the secondary isopimar-8(14)-en-15-yl⁺ that is strongly ion paired to the pyrophosphate, again as a transient intermediate or transition state that is quickly rearranged to the more stable tertiary abiet-8(14)-en-13-yl⁺, which is still adjacent to the pyrophosphate counter-ion, prior to concluding deprotonation. Notably, the production of isopimara-7,15-diene seems to result from use of a pre-existing active site base, rather than the introduced hydroxyl group, as the necessary deprotonation at C7 also occurs during abietadiene production, consistent with our mechanistic hypothesis (i.e. that the hydroxyl acts electrostatically rather than as a general base). In any case, the shift in active site location of this switch residue presumably reflects the difference in configuration between the enantiomeric (ent-CPP) substrate of the ent-(iso)kaurene synthases, relative to that (CPP) for AS.

Application
Implicit in the mechanistic hypothesis presented above is that the identity of the residue at this position acts as a true “switch”, including the ability to increase reaction complexity in addition to short-circuiting more complex reactions. This was investigated with a syn-pimara-7,15-diene synthase from rice, whose encoding gene location in a diterpenoid biosynthetic gene cluster\textsuperscript{35} dedicated to production of the derived momilactones\textsuperscript{55} indicates a long standing role in straight-forward tricycle production, and which contains a Thr at the previously identified ent-(iso)kaurene synthase product outcome switch position, consistent with its relatively simple reaction mechanism (i.e. deprotonation of the syn-pimar-15- en-8-yl\textsuperscript{+} intermediate formed by initial cyclization of syn- CPP). Upon substitution of this Thr with Ile, we found that the resulting mutant enzyme now largely produced a novel diterpene that was identified as syn-aphidicol-15-ene, whose production requires further cyclization and rearrangement of the initially formed syn-pimar-15-en-8-yl\textsuperscript{+} (Figure 8). Notably, previous biomimetic studies with syn-copalol had demonstrated that similar extended cyclization and rearrangements can occur in organic solvent,\textsuperscript{56} consistent with our observations that the presence of an aliphatic residue allows more complex reactions. However, it should be noted that the mutant enzyme more specifically catalyzes such extended cyclization-rearrangement, which presumably reflects, in part, carbocation migration towards the retained pyrophosphate anion co-product, as well as enabling substrate folding. Accordingly, our results demonstrate that the identity of the residue at this position can act as a true switch for diterpene synthase product outcome, and are consistent with our hypothesis for the underlying mechanism.\textsuperscript{54} In addition, the ability of substitution at the same position as in ent-(iso)kaurene synthases to alter product outcome in this syn-CPP reactive diterpene
synthase suggests that the determinant for the location of the switch residue must be the C-9 stereochemistry shared between ent- and syn-CPP (c.f. Figures 5 and 8).

However, it must be noted that our results with the rice syn-pimaradiene synthase appear to have been somewhat fortuitous. In particular, it is clear that other, presumably secondary changes can mask the ability of the pyrophosphate anion co-product to steer carbocation migration, as it takes additional active site residue substitutions to convert isopimaradiene synthase to the production of abietadienes. Notably, this also is consistent with our hypothesis that the hydroxyl group does not act as a general base, but rather exerts electrostatic effects on the catalyzed reaction.

**Stereospecificity**

As explicitly stated in our mechanistic hypothesis, the switch residue should be proximal to the initially formed pimar-15-en-8-yl carbocation, with its location varying with substrate stereochemistry. Consistent with this supposition, substitution of the ent-(iso)kaurene synthase switch residue in the normal CPP specific AS (Val→Thr) does not affect product outcome. More direct proof for proximity of the switch residue to the pimar-15-en-8-yl carbocation may be found in our recent observation that the introduction of Cys at the relevant switch position in AS results in a self-inactivating mutant enzyme. While further characterization is required, it is tempting to speculate that this derives from alkylation of the Cys side chain by pimar-15-en-8-yl\(^+\), which would directly demonstrate the hypothesized proximity. Regardless, alkylation of an introduced thiol (Cys), but not hydroxyl (Ser), is further consistent with our mechanistic hypothesis that the observed effect on product outcome is electrostatic in nature (i.e. rather than the hydroxyl acting as a general base).
Conclusions

Here we have discussed the recent findings from our group demonstrating the presence of a single residue switch for product outcome in the diterpene synthases relevant to labdane-related diterpenoid natural products biosynthesis, which is the central focus of our studies. While no directly analogous results have been reported with other terpene synthases, we suggest that our hypothesis for the underlying mechanism is more widely applicable. In particular, that the retained pyrophosphate anion co-product in the active site will steer carbocation migration towards itself in the absence of other countering electrostatic effects. The example arising from our results is the ability of a proximal hydroxyl dipole to stabilize the pimar-15-en-8-yl carbocation formed by initial cyclization of CPP, enabling terminating deprotonation, while the lack of such stabilizing interaction (i.e. when an aliphatic side chain is present instead) allows extended, more complex cyclization-rearrangement reactions, which seems to result from carbocation migration towards the pyrophosphate anion co-product. Such favorable ion pairing alleviates the accompanying, energetically unfavorable tertiary to secondary carbocation transition, as indicated by the strong potentiating affect of inorganic pyrophosphate on the corresponding aza-analogs. In addition, in each case the relevant secondary carbocation appears to be a high energy transition state that is rapidly rearranged to a tertiary carbocation that remains proximal to the pyrophosphate anion co-product, retaining some ion pairing potential. From consideration of the generalized terpene cyclization reaction, with intramolecular S_N' addition from a distal double bond on the incipient allylic carbocation arising from ionization of the diphosphate ester, the resulting pyrophosphate anion co-product will then be proximal to the original allylic double bond,
but distal from the initially formed carbocation. Hence, secondary addition from the pyrophosphate proximal double bond will form a carbocation that will be stabilized by ion pairing with the pyrophosphate anion. For example, examination of sandaracopimaradiene docked into the recently determined crystal structure of AS, along with diphosphate co-product, indicates that such carbocation migration from C8 to C15 brings the cation ~2.4Å closer to the diphosphate anion (5.4Å → 3.0Å; Figure 7). The counter-intuitive implication of this mechanism is that production of more “simply” cyclized terpenes requires electrostatic stabilization of the corresponding “final” carbocation by the enzyme, while the production of more “complex” products may not require any specific interaction with the enzyme beyond sterically imposed substrate folding. Consistent with broader applicability of this mechanistic hypothesis are recent theoretical studies indicating a role for the pyrophosphate anion co-product in the reaction catalyzed by monoterpenyl cyclases such as bornyl diphosphate synthase. 

Notes and references

"Department of Biochemistry, Biophysics, & Molecular Biology, Iowa State University, Ames, IA 50011, USA. *Fax: +1-515-294-0453; Tel: +1-515-294-8580; E-mail: rjpeters@iastate.edu"

† This article is part of the ‘Enzymes and Proteins’ web-theme issue for ChemComm.
‡ Financial support for the work from the authors group described here was provided by a grant from the NIH (GM076324) to R.J.P.

References

60. Y. J. Hong and D. J. Tantillo, Org Biomol Chem, 2010, 8, 4589-4600.
Figures

Figure 1: Binding of trinuclear cluster of divalent metal ions by conserved residues and pyrophosphate in bornyl diphosphate synthase.\textsuperscript{7}

Figure 2: Cyclization of FPP to germacrene A.
Figure 3: BPS catalyzed cyclization of GPP to bornyl diphosphate, with aza-analogs of early and late carbocation intermediates shown with relative orientation to that of the fixed inorganic diphosphate group.  

![Chemical Structure Image](image)

Figure 4: Protonation-initiated cyclization of GGPP to CPP stereo-isomeric intermediates in labdane-related diterpenoid biosynthesis.  

![Chemical Structure Image](image)
Figure 5: Cyclization and rearrangement of ent-CPP to ent-(iso)kaurene can be short-circuited to produce ent-pimaradiene by substitution of a specific Ile with Thr.\textsuperscript{47}

Figure 6: Cyclization and rearrangement of CPP to abietadiene can be short-circuited to produce isopimaradiene by substitution of a specific Ala with Ser.\textsuperscript{48}
Figure 7: Active site of abietadiene synthase with docked sandaracopimaradiene and inorganic diphosphate with trivalent cluster of magnesium ions. Also shown are distances to key C8 and C15 carbocation locations discussed in text, and the switch residue (Ala).

Figure 8: Cyclization-rearrangement of syn-CPP to syn-aphidicolene catalyzed by syn-pimaradiene synthase upon substitution of a particular Thr with Ile.\textsuperscript{54}
Chapter VII: A secondary position involved in catalytic specificity of labdane-related diterpene synthase

Manuscript in preparation and to be submitted to Phytochemistry

Ke Zhou and Reuben J. Peters*

Abstract

A single amino acid residue switch on product outcome in plant labdane-related diterpene synthases has been reported. However, compared with the conversion “short circuiting” the complex cyclization and rearrangement reaction, such as from a specific isokaurene synthase into a specific pimaradiene synthase, the reverse conversion, which increases complexity of diterpene synthase reaction, often decreases product specificity of resulting enzymes. We further investigated a nearby valine/leucine residue which also seems to be involved in product specificity as we found that increasing side chain volume decreases the product specificity of rice isokaurene synthases, leading to increased yield of side products. These results highlight the importance of the secondary position in the active site to product outcome.

Introduction

Terpene synthases catalyze the most complex chemical cyclization and/or rearrangement reactions occurring in nature, and are of significant mechanistic interest, especially in the mechanisms by which they exert control over the catalyzed carbocationic reaction. With such ability, they create the most structurally manyfold hydrocarbon backbones which underlie the most diverse class of natural products, terpenoids, with over 50,000 already known.

Recent studies have demonstrated that terpene synthases exhibit extreme plasticity, with small numbers of amino acid substitutions being sufficient to significantly alter product
outcome (Xu et al. 2007b; Wilderman et al. 2007; Morrone et al. 2008; Keeling et al. 2008; Yoshikuni et al. 2006; Kampranis et al. 2007). Most strikingly, the ability of single residue changes to dramatically shift the product outcome of labdane-related terpene synthases has already been successfully demonstrated (Xu et al. 2007b; Wilderman et al. 2007; Morrone et al. 2008).

Labdane-related diterpene synthases is served as a model system for analysis of terpene synthase substrate and product specificity. Usually, class II diterpene synthase cyclizes GGPP by a protonation-initiated reaction to a bicyclic intermediate, labdadienyl/copalyl diphosphate (CPP). This biosynthetic intermediate is then further cyclized and/or rearranged by a separate ionization-initiated (class I) terpene synthase, or separate domain at same enzyme (bifunctional), leading to a special large group of natural products (~7,000 known) termed the labdane-related diterpenoids (Peters 2006).

The functional genomics investigation of rice (Oryza sativa) led to identification of highly homologous functionally divergent paralogs (Xu et al. 2007a). OsKSL5i (indica subspecies) produces tetracyclic ent-isokaur-15-ene, while OsKSL5j (japonica subspecies) makes tricyclic ent- pimara-8(14),15-diene, representing direct deprotonation of the proposed ent-pimar-15-en-8-yl+ intermediate in the cyclization of ent-CPP to kaurene type diterpenes. This pair of orthologs shares 98% identity at the amino acid sequence level, and there are only three differences in the modeled active site. So a single residue switch was identified to contribute their functional difference, alteration mutually of which was sufficient to change product outcome, with the presence of a Thr resulting in “short circuiting” reaction leading to production of ent-pimaradiene, while an Ile leads to more complex cyclization and rearrangement reaction to produce ent-isokaurene. Strikingly, in two disparate ent-kaurene
synthases (OsKS1 and AtKS), Thr substitution at the corresponding position had a very similar effect, “short-circuiting” the more complex cyclization leading to ent-kaurene, instead producing the “simpler” ent-pimaradiene (Xu et al. 2007b). In addition, analogous functionally coupled variation was found in labdane-related diterpene synthases acting on CPP of normal and syn- stereochemistry (Wilderman et al. 2007; Morrone et al. 2008). Although the corresponding position of single residue switch does not exhibit the same at normal CPP specific diterpene synthase, in comparisons of those at ent- or syn- CPP specific diterpene synthases, it has been hypothesized in all of these reports that the side chain of this product switch residue is competed with the tightly bound pyrophosphate anion co-product to control the fate of carbocation intermediates (Xu et al. 2007b; Wilderman et al. 2007; Morrone et al. 2008).

However, compared with the OsKSL5i:I664T which converted the specific ent-isokaurene synthase to the specific ent-pimaradiene synthase, OsKSL5j:T664I no longer produced pimaradiene, but rather a mixture of isokaurene, atiserene and kaurene (50:37:13). So OsKSL5j:T664I exhibited decreased product specificity compared with wild type OsKSL5i (Xu et al. 2007b). Such disparity prompted the further investigation in the effect on product outcome from the nearby residues in the active.

**Results**

Previous mutational analysis demonstrated the critical single residue switch related to the product specificity in plant diterpene synthases (Xu et al. 2007b). However, the OsKSL5i:I664T can convert a specific isokaurene synthase into a specific pimaradiene synthase, while, a reverse mutant, OsKSL5j:T664I no longer produced pimaradiene, but rather a mixture of isokaurene, atiserene and kaurene (50:37:13) (Figure 1). So
OsKSL5j:T664I exhibited decreased product specificity in comparison with wild type OsKSL5i (above 90% isokaurene, Figure 1) (Xu et al. 2007b). To figure out if one or more of other residues difference between OsKSL5i and OsKSL5j could affect outcome profile, other two residues in the active site was tested. Our results demonstrated that OsKSL5j:L661V/T664I produced more isokaurene to 81% at the expense of atiserene and kaurene (Figure 1). It seems that the shorter aliphatic side chain makes a more specific isokaurene synthase. The Valine substitution for Leucine at the corresponding position of OsKSL5i (V661L) decreased its catalytic specificity, with production of only less than 70% isokaurene and more atiserene and kaurene, indicating the importance of this position for later rearrangement and deprotonation reactions. A similar observation on OsKSL 6 also confirmed that. An introduced Leu instead of Val resulted in a slightly promiscuous mutant, with the increase of atiserene to 5%, kaurene to 7%, in comparison with none of them detected in the wild type enzyme (Figure 1). However, this effect did not extent well to kaurene synthases, like AtKS.

**Discussion**

(Di)terpene synthases catalyze the remarkably complex chemical reactions and exhibit significant plasticity. The template model has been proposed and well received to generally describe the control these enzymes exert on the reaction trajectory. Specifically, it implies that terpene synthases dictate reaction outcome by restricting their substrate and discrete reaction intermediates to a subset of the possible conformations (Christianson, 2008). However, recent results highlighted electrostatic effects also exerted by these enzymes. This is emphasized by the striking effects of a recently uncovered single residue switch for
labdane-related diterpene synthase product outcome (Xu et al. 2007b; Wilderman et al. 2007; Morrone et al. 2008).

Interestingly, although aliphatic hydroxyl residue substitution for aliphatic residue, such as OsKSL5i:1664T, presented a complete conversion from the specific isokaurene synthase into a specific pimaradiene synthase, converse change needed at least one more nearby residue substitution to overcome promiscuous production. Actually, regardless of the proportion of expected isokaurene, all these three products including isokaurene, kaurene and atiserene share a similar catalytic pathway from start through the ent-beyer-15-yl+ intermediate. Therefore, it does not weaken the previous hypothesis that primary single residue switch is adjacent to the ent-pimar-15-en-8-yl carbocation, such that the presence of a hydroxyl dipole stabilizes this intermediate long enough for deprotonation to occur, leading to “short-circuit” complex cyclization and/or rearrangement reactions, while the lack of such stabilization (i.e. in the presence of an aliphatic residue) enables the tightly bound pyrophosphate anion co-product to steer carbocation migration towards itself, with subsequent rearrangement to the more stable tertiary ent-kauran-16-yl+, followed by concluding deprotonation (Zhou & Peters 2011). The secondary position we are discussing here mostly seems to direct the latter steps including ring arrangement by which separates ent-atiserene from kaurene-like structure, as well as final deprotonation at the different position leading to ent-kaurene or ent-isokaurene. Indeed, it sometimes takes additional active site residue substitutions to convert one from producing simple pimaradiene structure to specifically synthesizing a compound with increased complex. In the case of this pair of highly homologous paralogs, at least, the ability of the pyrophosphate anion co-product to steer carbocation migration is absolute when the hydroxyl residue at primary position is substituted by aliphatic one, while a trial of
similar single substitution failed to convert isopimaradiene synthase to the production of abietadienes without the additional active site residue substitution (Keeling et al., 2008).

Thus, the results of current secondary residue together with those of previous uncovered primary single residue switch indicates the ability of terpene synthase to exert combination roles of template mode and electrostatic effects on product outcome. However, on the basis of understanding the permissiveness of the active site contour that assists the orientations and conformations of flexible substrate and intermediate during cyclization reactions, resulting in the fidelity and promiscuity in product profile (Christianson 2008), it is hard to explain how the shorter side chain (V to L) makes less permissive contour without precise structure in active site of relevant diterpene synthase.

**Conclusion**

The investigations on the secondary valine/leucine residue nearby the primary single residue switch in relevant labdane-related diterpene synthases demonstrated its effect on product specificity, and supported the importance of active site contour to orient specific intermediates. Combined with electrostatic effects exerted by single residue switch, controlling over catalyzed carbocationic reaction by means of two approaches seems to keep respective responsibility well, at least for labdane-related diterpene synthases.

**Experimental**

**General Procedures**

Authentic standards for the diterpenes identified here were kindly provided by Robert Coates (University of Illinois at Urbana–Champaign, Urbana, IL). Unless otherwise noted, all other chemicals were purchased from Fisher Scientific (Loughborough, Leicestershire, U.K.), and molecular biology reagents were from Invitrogen (Carlsbad, CA). GC was
performed with Agilent (Palo Alto, CA) 6890N GC instruments by using an HP-1 column with FID or an HP-5 column with MS detection using a 5973N mass-selective detector in electron ionization (70 eV) mode located in the W. M. Keck Metabolomics Research Laboratory at Iowa State University. Methods for running the samples are the same as previous description (Xu et al. 2007b).

**Recombinant Constructs**

The OsKSL5i and OsKSL6 genes have been previously described (Xu et al. 2007a), whereas AtKS was previously cloned and kindly provided by Shinjiro Yamaguchi (Institute of Physical and Chemical Research, Wako, Japan) (Yamaguchi et al.). OsKSL5j was also previously cloned (Xu et al. 2007b). All genes were transferred into the Gateway vector system (pENTR). Sitedirected mutagenesis was carried out via PCR amplification of the pENTR constructs with overlapping mutagenic primers, and the mutant genes were verified by complete sequencing. The resulting wild-type and mutant genes were then transferred via directional recombination to the T7-promoter N-terminal GST fusion expression vector pDEST15.

**Enzymatic Analysis via metabolic engineering**

Functional characterization of the all wild type and mutants were accomplished by use of our previously described modular metabolic engineering system (Cyr et al. 2007). Briefly, class I labdane-related diterpene synthases such as OsKSL in pDEST15 expression vectors were co-expressed with pACYC-Duet (Novagen/EMD) derived plasmids (pGGeC), where GGPP and ent stereospecific CPP synthases were carried on. Subsequently, C41 cells were co-transformed with pDEST expression vectors carrying OsKSL genes (WT or mutants), along with pGGeC vectors. Growing of the recombinant C41 cells followed the procedure
like previously described (Cyr et al. 2007). The resulting cultures were then extracted by hexane with the same volume. The diterpene products were identified by comparison to authentic standards via GC-MS, and relative amounts were quantified by GC-FID.

**Acknowledgment**

We thank Prof. Robert M. Coates for providing authentic standards, and Dr. Shinjiro Yamaguchi for the AtKS clone. This work was supported by National Institutes of Health Grant GM076324; National Science Foundation Grant MCB-0416948; and U.S. Department of Agriculture Cooperative State Research, Education, and Extension Service National Research Initiative Grant 2005-35318-15477 (to R.J.P.).

**References**


Figures

Figure 1: Identification of products formed by OsKSL5i, OsKSL5j, OsKSL6, and their mutants via GC-MS based comparison to authentic standards. (a-g) Chromatograph of the products formed by OsKSL5i wt, OsKSL5i-V661L, OsKSL5j wt, OsKSL5j-T664I, OsKSL5j-T664I/L661V, OsKSL6 wt, and OsKSL6-V661L, respectively, from ent-CPP.
Chapter VIII: Conclusion and future direction

The structure of the representative bifunctional plant diTPS, AgAS, presented here provides insights into both the enzymatic mechanism and evolution of TPS more generally. The overall arrangement of the γβα domain structure provides a rationale for the observed βα domains structure of most plant lower TPS. Indeed, not only does its βα domain exhibit expected very high structural homology with other plant lower class I TPS, but, in particular, the observed structural homology of the class II active site containing γβ domains with triterpene synthases is consistent with homologous origins for these mechanistically similar enzymes.

In the class II active site, it is now possible to visualize the arrangement of catalytic residues, including not only the previously noted D_{402}XDD motif (Prisic et al. 2007), but other neighboring residues as well. For example, N451, which is conserved in all plant class II TPS, forms a hydrogen bond with the catalytic acid D404, presumably to increase its ability to carry out the energetically unfavorable protonation of a carbon-carbon double bond. In addition, another completely conserved tryptophan, which faces the active site and plays an important role in catalysis (Peters et al, 2002), seems likely to be involved in stabilizing carbocation intermediates via interactions with its π quadrupole. Intriguingly, close to this aromatic residue is a residue that previous results indicated controls the susceptibility of class II diterpene cyclases to Mg^{2+}-dependent inhibition (Mann et al. 2010). However, the corresponding R356 in AgAS is found on the opposite side of the relevant helix and reaches to the protein surface rather than forming any part of the active site, not supporting our original hypothesis that this residue, conserved as a histidine or arginine in all plant class II diterpene cyclases, would be directly involved in catalysis (Mann et al. 2010). Thus, we have
suggested that the mechanism underlying this remarkable regulatory effect on catalysis seems likely also to involve the neighboring aromatic residues, although this remains unclear as of yet. Also opaque, unfortunately, is the role of Mg$^{2+}$ in class II TPS due to their absence in our structure, as well as that of AtCPS. Thus, the mechanisms whereby Mg$^{2+}$ exerts both activating and inhibitory effects remains unclear. However, it will be necessary to elucidate these. Since current structures of AgAS and AtCPS lack Mg$^{2+}$, and that of AtCPS co-crystallized with the 15-aza-GGSPP without Mg$^{2+}$ does not display the expected substrate orientation for cyclization (Köksal et al. 2011), locating Mg$^{2+}$ in both activating and inhibitory binding sites in the class II TPS would provide further insight into the relevant mechanisms. One way possible to achieve this is to soak Mg$^{2+}$, in the presence or absence of substrate analog, such as 15-aza-GGSPP, or product analog, like 2-fluoro-CPP, respectively, into native crystals. Co-crystallization at wider crystallization conditions could be another method, although, unfortunately, such efforts have not yet yielded high quality crystals to date. Ideally, by anchoring the diphosphate moiety, activating Mg$^{2+}$ binding can impose correct orientation of the GGPP analog for cyclization. Such structures might also clarify how Mg$^{2+}$ at higher concentrations exerts an inhibitory effect on catalysis, at least in those CPS involved in gibberellin biosynthesis. In addition, given the production of different stereoisomers of CPP by AgAS and AtCPS, co-crystallized structures with correctly oriented and folded GGPP analogs, or their respective CPP analog (normal or syn), could expose critical structural elements underlying product stereochemistry.

Interestingly, the density map for our AgAS crystal structure revealed a mysterious hydrophobic molecule bound in the class II active site pocket. It is unclear yet what this, as it can not be identified with the current resolution, but all components of crystallization
conditions, protein storage buffer, and cryo-protectant can be excluded from fitting into that density well, suggesting this may have been bound into this active site before crystallization. Nevertheless, determination of its structure could provide insight into specific recognition on the structural level and enable design of new inhibitor against class II diterpene synthases.

In the class I active site, besides the characteristic DDXXD and (N,D)DXX(S,T,G)XXXE metal binding motifs, our AgAS structure further enables visualization of the previously identified single residue switch for product outcome (Wilderman et al. 2007), demonstrating that this side chain points into the active site. A sandaracopimara-8(14),15-diene docked into the AgAS class I active site further indicates proximity of the Ala side chain to carbon 8, supporting a previous hypothesis that this side chain is in close proximity to the carbocation in the isopimar-15-en-8-y1+ intermediate formed by initial cyclization. This is further consistent with our mechanistic hypothesis that substitution with a hydroxyl containing Ser stabilizes the initially formed carbocation at this position sufficiently long enough for deprotonation to occur, yielding the mixture of isopimara-7,15-diene and sandaracopimara-8(14),15-diene observed as the products of the corresponding AgAS:A723S mutant (Wilderman et al. 2007). However, it is not unreasonable that the docking result could be questioned. Our current AgAS structure does not have substrate or product analogs bound in the class I active site. Indeed, it appears to be in an open conformation for this active site according to the well-studied homologous structures of lower terpene synthases (Starks et al. 1997; Rynkiewicz et al. 2001; Whittington et al. 2002). Thus, it seems well worth trying co-crystallization with a substrate analog, such as 2-fluoro-CPP, and cofactor Mg$^{2+}$, or with a reaction intermediate analog, such 15-aza-pimarenly, Mg$^{2+}$, as well as inorganic pyrophosphate co-product, to provide visualization of the expected proximity between not
only carbon 8 and residue switch (A723), but also the aza group and pyrophosphate in the latter case.

Although several informative structures of class I TPS, as well as successful studies via mutagenesis, commonly demonstrate that the substrate conformation imposed by the TPS active site plays a critical role in the ability of these enzymes to direct specific product outcome (Christianson 2006; Christianson 2008), it is not well understanding how labdane-related diterpene synthases distinguish substrates with different stereochemistry, nor how TPS more generally control high reactive carbocation intermediates. Our stable of labdane-related class I diterpene synthases may provide suitable targets for more crystallization trials. Success on any of them which utilize syn-, ent-, or any pair of isomers of CPP as substrate(s) would provide the ability to compare their active site contours with that of AgAS, ideally leading to insight for structural determinants for substrate recognition. Redesigning the active site cavity of AgAS, especially around desired carbocation intermediate conformations, to enforce alternative substrate conformations also may generate enzymes with new product profiles. Such results would provide deeper understanding for the role of substrate folding, and other factors directing the catalyzed reaction cascades, and enable design of novel enzymatic function.

Another aspect of enzymatic catalysis explored in this thesis is further study on the specificity of the diterpene synthases relevant to labdane-related diterpenoid natural products biosynthesis, following the recent findings from our group demonstrating the presence of a single residue switch for product outcome in this class of terpene synthase (Xu et al. 2007; Wilderman et al. 2007; Morrone et al. 2008). In conclusion, we have demonstrated here that the single residue switch for product outcome in labdane-related diterpene synthases is
specific with respect to both active site location (AgAS:A723S versus AgAS:V727T) and side chain chemical composition (AgAS:A723S versus AgAS:A723C). Furthermore, our results support the previously implied mechanism wherein the side chain of the residue at the switch position directly interacts with the initially formed pimar-15-en-8-yl carbocation. A polar group, particular more electronegative hydroxyl oxygen might provide the transient stabilization and thus sufficient time for deprotonation of that intermediate to occur, leading to ‘short circuiting’ of more complex cyclization and rearrangement reactions. Measurement by quadrupole time-of-flight (Q-Tof) mass spectrum (MS) and trypsin digestion followed by MS demonstrated that a mass shift difference of 272 Da between AgAS:A723S samples incubated in the presence and absence of GGPP, indicating geranylgeranylation of this mutant during catalysis, and proximity of at least one carbocation intermediate to 723 position. Basically, if we can figure out the structure of the captured molecule, we could demonstrate which carbocation intermediate is captured by alkylation, presumably by the thiolate of the introduced Cys. Typically, structure determination of this small molecule could be accomplished via comparison to authentic standards, or solved by NMR. It is necessary for both approaches to digest the alkylated AgAS:A723C to a mixture of amino acids firstly. For the first method, synthesis of a couple of desired authentic standards could be viable, while for the other approach, separation and detection of the target molecule is also difficult due to its low abundance. Crystallization of alkylated AgAS:A723C is another viable strategy to investigate this issue.

In this thesis, seven members of the wheat ent-kaurene synthase-like gene family (TaKSLs), as well as a putative ent-kaurene synthase from barley (HvKS) were cloned and biochemically characterized. The HvKS was verified as encoding an ent-kaurene synthase.
TaKSL6 also produces ent-kaurene exclusively, and both were considered to function in gibberellin biosynthesis. Although another pair of genes, TaKSL5-1 and 5-2, encoded class I diTPS that produce ent-kaurene as their major product, these also produce substantial amounts of ent-beyrene, and lacked the insertional domain found in all other plant KS. Such rare arrangement in plant diterpene synthases is actually occurring consistently in lower terpene synthases (i.e. sesquiterpene or monoterpene synthases). Wheat contains the inducible TaKS11, which encodes an enzyme producing isopimaradiene, which is probably decorated by downstream oxidase to form phytoalexins, while the uninducible TaKS14, which encodes a KSL that produces pimara-8(9)-15-diene, may be involved in metabolism of phytoanticipins or allelochemicals. Of particular interest, these two KSL can react with the syn-CPP that does not seem to be made in wheat and, together with the dually reactive TaKS2, demonstrate substrate plasticity similar to that found in the closely related cereal crop plant rice (Morrone et al. 2011), which suggests that such plasticity plays a role in diterpenoid metabolic evolution.

References


